Theor Chim Acta (1985) 68, 363-377

THEORETICA CHIMICA ACTA

A note on the attainability of states by equalizing processes

C. Zylka

Sektion Physik, Karl-Marx-Universität, DDR-7010-Leipzig, German Democratic Republic
(Received February 4, revised April 26/Accepted June 10, 1985)

In this paper, we consider "regions of attainability" in certain state spaces as a function of the initial state under a well-defined and physically relevant class of processes. These processes are the continuous and "chaos-enhancing" processes (in the sense of Ruch and Uhlmann). It turns out that these sets have complicated geometrical structures: They are polyhedra which are generally non-convex. The proof - a rather geometrical one - is given. A theorem of Hardy, Littlewood, and Polya and properties of bistochastic matrices are used crucially.

Key words: Order structure of states — The half-order "$>$" — Chaos-enhancing processes — Regions of attainability

1. Introduction

1.1. General remarks

This investigation is based on the concept of the *order structure of states*, which was put forward independently by Ruch and Schönhofer [8], and Uhlmann [9] in the early 70's. Whereas Ruch and Schönhofer considered classical distributions following their concept of "measuring the extent of identification", Uhlmann pursued his idea that there should be a finer distinction for (quantum) states than that of merely being pure or mixed. The result was that it seemed reasonable to introduce a partial order within very general state spaces, symbolically: "$>$", in words: "more mixed than".

There are many hints (esp. from the quantum case) that this concept should be a useful frame to understand more about irreversibility on a rather fundamental level. For a recent survey on this line, see, e.g. Alberti and Uhlmann [2, 3].

As is known, *dissipative motion* shows the following property: coming from the boundary of the state space into the interior, the states "lose structure" (or purity, or information). This feature of dissipative systems can be reflected by the partial order very clearly. Therefore soon after introducing this relation, one started to study *processes* which are defined just by the property that every later state is more mixed than every earlier one. These processes are called *mixing enhancing processes* (or *c-processes*).

It is also known that a lot of processes which appear in physics or chemistry are of this kind. We only mention examples of the type

(i) $d/dt\, p^i = \sum_k L_{ik} p^k$

and

(ii) $d/dt\, p^i = \sum_{j,k,l} (A_{ijkl} p^k p^l - A_{klij} p^i p^j)$.

The former is the well-known master equation, the latter is the Boltzmann-Carlemann equation (see Chap. 5).

In this paper, we won't study *c-processes* in detail but under the aspect of the *attainability of states*. By this term we abbreviate the problem whether or not states are attainable from a fixed initial state by a certain process. This question will be answered for continuous c-processes over classical discrete states, i.e. for finite dimensional probability vectors. We are able to provide an explicit geometrical construction for their sets of attainability. Moreover, we can prove essential properties of them: they are polyhedra which turn out to be not convex, in general. All proofs use elementary tools like results on double stochastic matrices, convexity arguments, and topology.

The following parts of the introduction provide a simple physical model which is convenient for the construction procedure and a survey of the notations and definitions we will use. The second chapter presents the origin of the problem and earlier contributions to its solution. In the third chapter the sets of attainability for heat conduction processes will be constructed. In the following chapter we will show that these sets are identical with the sets which correspond to the much more general continuous c-processes. The fifth chapter closes the paper with some discussion.

1.2. Our model system

In order to illustrate the problem we now introduce a special physical system.

Let us consider n (n finite) bodies with equal heat capacity. Each of these bodies should be in thermal equilibrium with itself and therefore characterized by just one temperature. This temperature is the only property of the body we are interested in.

If the bodies are numbered (arbitrarily, but once chosen then fixed forever), then the state of our system is completely described by its temperature distribution.

Next we consider a "mechanism" in order to change states: For this purpose we assume an "ideal wire". By this wire we connect any two bodies for any period of time. An "amount of heat" will be transported from the hotter body to the cooler one. This procedure is called a "two-body heat exchange". Then we connect two other bodies and so on.

Later - and this is the essential result - we will replace this special physical mechanism by a much more general, but also well characterized class of processes, the so-called "continuous c-processes".

The problem we solved is: give a description of the set of all temperature distributions (i.e. all states) attainable by these processes starting at an arbitrary initial state.

1.3. Notations, definitions etc.

To deal with the above problem we provide some tool: Let the bodies be numbered by $i \in \{1, 2, \ldots, n\} := \Omega$, then p^i denotes the temperature of the ith body. The temperatures of all bodies are collected in a vector $\boldsymbol{p} := (p^1, p^2, \ldots, p^n)$. According to the first law of thermodynamics their sum is constant and will be - without loss of generality - normalized to 1. Therefore we can consider the $\boldsymbol{p}$ as probability vectors. Thus the whole state space is:

$$\mathcal{S}_n := \left\{ \boldsymbol{p} : \boldsymbol{p} = (p^1, p^2, \ldots, p^n),\ p^i \geq 0, \sum_i p^i = 1 \right\}.$$

We further recall the so-called "bistochastic" matrices:

$$\mathrm{BST}_n := \left\{ T\colon T \text{ is a } (n \times n)\text{-matrix},\ T_{ik} \geq 0,\ \forall i, k; \sum_i T_{ik} = \sum_k T_{ik} = 1 \right\}.$$

One can see that these matrices are stochastic ones with the additional property of keeping the equipartition $\boldsymbol{e} := (1/n, 1/n, \ldots, 1/n)$ invariant. Now we consider a subset of BST_n:

$K_{(2)n} := \{T\colon T \in \mathrm{BST}_n$ and has the structure

$$T = \begin{array}{cc}
\begin{array}{ccccccccccc} & & & k & & & & l & & & \end{array} & \\
\left(\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
& \ddots & & & & & & & & & \\
& & 1 & & & & & & & & \\
& & & a & & & & 1-a & & & \\
& & & & 1 & & & & & & \\
& & & & & \ddots & & & & & \\
& & & & & & 1 & & & & \\
& & & 1-a & & & & a & & & \\
& & & & & & & & 1 & & \\
& & & & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{array}\right) &
\begin{array}{c} \\ \\ \\ k \\ \\ \\ \\ l \\ \\ \\ \\ \end{array}
\end{array}$$

with $a \in [\frac{1}{2}, 1]$ and $1 \leq k < l \leq n$, all other matrix elements are equal to 0}.

The reader should realize that these matrices applied to a state effect just such two-body heat exchanges as explained in Sect. 1.2. Therefore we are now able to give a precise formulation of the set of those states attainable by a succession of these exchanges, namely

$$K^w(\boldsymbol{p}) := \left\{\boldsymbol{q} \colon \boldsymbol{q} = T\boldsymbol{p} \text{ with } T = \prod_i T_i,\ T_i \in K_{(2)n},\ \boldsymbol{p}, \boldsymbol{q} \in \mathcal{S}_n,\ \boldsymbol{p} \text{ fixed}\right\}.$$

The "w" stands for "Wärme"; for short we call a product $\prod_i T_i$ with $T_i \in K_{(2)n}$ applied to a state *a* "w-process", see also Sect. 4. We take this occasion to fix a result which follows immediately from the definition of $K^w(\boldsymbol{p})$:

Lemma 1. *Let $\boldsymbol{q} \in K^w(\boldsymbol{p})$ and $\boldsymbol{q} \neq \boldsymbol{p}$. Then $K^w(\boldsymbol{q}) \subsetneq K^w(\boldsymbol{p})$. (Iff $\boldsymbol{q} = \boldsymbol{p}$ the equality sign is true.)*

Now we are going to define a relation between states, compare Uhlmann [2]: we say the state $\boldsymbol{q}$ is more mixed than the state $\boldsymbol{p}$ if there exists a $T \in \mathrm{BST}_n$, so that $\boldsymbol{q} = T\boldsymbol{p}$, symbolically $\boldsymbol{q} \succ \boldsymbol{p}$. (In another approach this is not a definition but a theorem by Hardy et al. [5]). Using this relation - which is a partial order - we set:

$$G(\boldsymbol{p}) := \{\boldsymbol{q} \colon \boldsymbol{q} \succ \boldsymbol{p},\ \boldsymbol{p}, \boldsymbol{q} \in \mathcal{S}_n,\ \boldsymbol{p} \text{ fixed}\}$$

Finally we recall a famous structure theorem:

Theorem 1 (Birkhoff). BST_n *is identical with the convex hull of the permutation matrices of dimension n.*

1.4. The geometrical aspect

Now we present a suitable representation of the state spaces. As one sees the cases $n = 1$ and $n = 2$ are not very interesting: For $n = 1$ nothing will happen and two bodies can equalize their temperatures, partially or totally. Therefore we will not mention the two cases throughout this paper and start at $n = 3$: $\mathcal{S}_3$ should be a regular triangle, its altitudes are the axes p^1, p^2 and p^3. Because of $\sum p^i = 1$

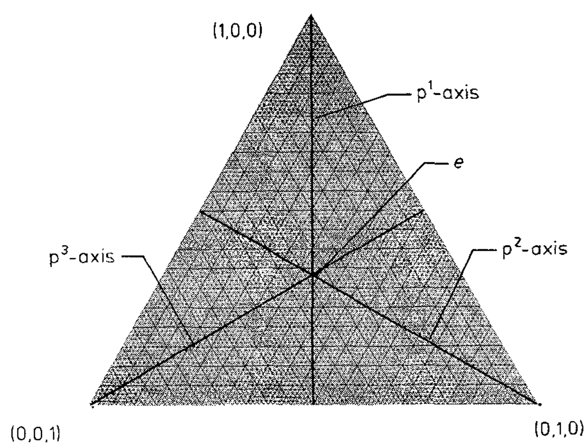

Fig. 1. The state space $\mathcal{S}_3$

the problem becomes a two-dimensional one, compare Fig. 1. The $\mathscr{S}_n$ with $n>3$ are introduced recurrently: $(n+1)$ pieces of $\mathscr{S}_n$ form the "boundary-hyperfaces" of $\mathscr{S}_{n+1}$. Thus $\mathscr{S}_4$ is a tetrahedron with analogous "tetrahedron coordinates", and so on.

Fig. 1 also shows that there is a natural division of $\mathscr{S}_n$ into $n!$ cells. This reflects - see the Birkhoff theorem - the permutation symmetry of the situation, because e.g. there are $n!$ possibilities to number the bodies.

A state the numbering of which fulfills $p^1 \geq p^2 \geq \cdots \geq p^n$ is called a canonical one and a state that has some p^i equal is called a symmetrical one.

Sometimes we also number the cells by $\varepsilon^j, j=1, 2, \ldots, n!$. The cell of the canonical state is called elementary cell. From now on it will be called ε^1 and the others remain unspecified. Finally we assume that the initial state is always a canonical one. It is clear that this is only a question of the numbering of the bodies.

2. What was known before?

The origin of our investigation lies in a mathematical result collected and newly proved by Hardy et al. [5]. They themselves used an old Muirhead [7] theorem of 1903. Equipped with the physical background explained in Sect. 1.2. one can also find it with Alberti and Uhlmann [2].

For convenience we present it in two equivalent formulations, the latter one is already adopted to our aim.

Theorem 2′ (Muirhead). *Let $\boldsymbol{p}$, $\boldsymbol{q} \in \mathscr{S}_n$ with $p^1 \geq p^2 \geq \cdots \geq p^n$ and $q^1 \geq q^2 \geq \cdots \geq q^n$. Then $G(\boldsymbol{p}) \ni \boldsymbol{q} \Leftrightarrow \boldsymbol{q} \in K^w(\boldsymbol{p})$.*

Proof ($\Rightarrow$). This can be found with Hardy et al. [5], Theorem 45, Lemma 2 (Attention: $\succ$ is used in the opposite direction !). ($\Leftarrow$): this is clear, because $K_{(2)n} \subset \mathrm{BST}_n$. $\square$

Theorem 2″. *Assumptions as in Theorem 2′. Then $G(\boldsymbol{p}) \cap \varepsilon^1 = K^w(\boldsymbol{p}) \cap \varepsilon^1$.*

Note that Theorem 2″ has an impressive verbal formulation: "In the elementary cell, $K^w(\boldsymbol{p})$ and $G(\boldsymbol{p})$ coincide".

Nevertheless it is also clear that $(K^w(\boldsymbol{p}) \cap \varepsilon^1) \subsetneq K^w(\boldsymbol{p})$ is valid, because the two-body exchanges will destroy the canonical order in general and the state will leave the elementary cell.

On the other hand it is true that: $K^w(\boldsymbol{p}) \subsetneq G(\boldsymbol{p})$, because e.g. the permuted states are in $G(\boldsymbol{p})$, but surely not in $K^w(\boldsymbol{p})$, since they are forbidden by the second law of thermodynamics.

Thus we have: $G(\boldsymbol{p}) \cap \varepsilon^1 \subsetneq K^w(\boldsymbol{p}) \subsetneq G(\boldsymbol{p})$, and therefore a rough hint where to look for the sets $K^w(\boldsymbol{p})$.

For $n=3$ the situation is illustrated in Fig. 2. (In order to construct $G(\boldsymbol{p})$ the Birkhoff theorem is used again.)

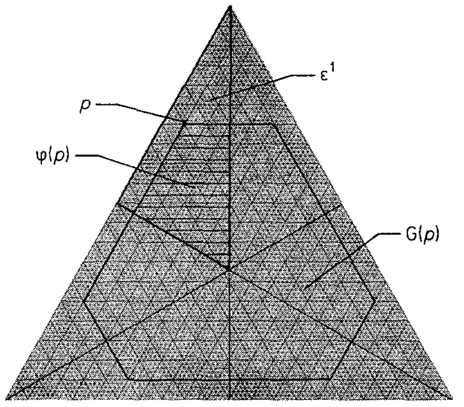

Fig. 2. An illustration of Muirhead's theorem

3. The first result: the sets $K(\boldsymbol{p})$

3.1. An overture

To get started the solution for $n=3$ is immediately given: The strongly framed area is $K^w(\boldsymbol{p})$, which turns out to be of nontrivial shape. The shaded area is that subset of it which was previously known - according to Theorem 2″ - and the hexagon of Fig. 2 is $G(\boldsymbol{p})$. This chapter will clarify that this solution is true and what can be generalized for higher dimensions.

For better readability we divide it into some steps.

3.2. An extension of Theorem 2″

It is possible to show that for an arbitrary $\boldsymbol{p} \in \mathcal{S}_n$, i.e. not necessarily out of ε^1, there exists a p-depending neighbourhood $E(\boldsymbol{p})$ such that $K^w(\boldsymbol{p})$ and $G(\boldsymbol{p})$ coincide in it. In other words, we want to replace the ε^1 of Theorem 2″ by convenient $E(\boldsymbol{p})$ and prove their existence by constructing them.

At first we define an additional ε^j, namely: $\varepsilon^0 := \overline{C\mathcal{S}_n} = \overline{R^{n-1} \backslash \mathcal{S}_n}$. Doing this we leave the physical region, but we gain mathematical uniformity. We note that all ε^j, $j=0, 1, \ldots, n!$ are closed sets with non-empty interiors.

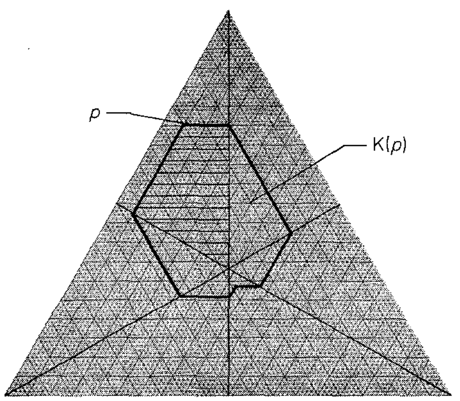

Fig. 3. The solution: $K(\boldsymbol{p})$ for $n=3$

It is not important that $\boldsymbol{p}$ and $\boldsymbol{q}$ are out of ε^1, it is sufficient for them to be from one and the same cell. By simultaneous denumbering we can trace them back to ε^1. Therefore we have:

Lemma 2. *Let $\boldsymbol{p}\in\varepsilon^j$, $j\in\{1,\dots,n!\}$. Then $K^w(\boldsymbol{p})\cap\varepsilon^j=G(\boldsymbol{p})\cap\varepsilon^j$.*

Now we consider symmetric states. Such states belong to more than one cell. We call $J(\boldsymbol{p}):=\{j:\boldsymbol{p}\in\varepsilon^j\}$. Lemma 2 is valid for each of these cells: $K^w(\boldsymbol{p})\cap\varepsilon^j=G(\boldsymbol{p})\cap\varepsilon^j$, $\forall j\in J(\boldsymbol{p})$. The union over these j gives $K^w(\boldsymbol{p})\cap E(\boldsymbol{p})=G(\boldsymbol{p})\cap E(\boldsymbol{p})$ with $E(\boldsymbol{p}):=\cup\,\varepsilon^j$, $\forall j\in J(\boldsymbol{p})$.

One can easily prove that $\boldsymbol{p}$ always lies in the interior of $E(\boldsymbol{p})$, i.e. $E(\boldsymbol{p})$ is, indeed, a neighbourhood of $\boldsymbol{p}$. Thus we arrive at:

Theorem 2‴. *For an arbitrary $\boldsymbol{p}\in\mathcal{S}_n$ the construction $\boldsymbol{p}\to E(\boldsymbol{p})$ gives a closed neighbourhood $E(\boldsymbol{p})$ with the property $K^w(\boldsymbol{p})\cap E(\boldsymbol{p})=G(\boldsymbol{p})\cap E(\boldsymbol{p})=:\varphi(\boldsymbol{p})$.*

3.3. An explicit description of $\varphi(\boldsymbol{p})$

Also the sets $\varphi(\boldsymbol{p})$ mentioned in Theorem 2‴ can be constructed explicitly[1]. On this occasion we introduce some "strategic" states, which are useful for later considerations.

Take a state $\boldsymbol{p}\in\mathcal{S}_n$ $(n\geq 3)$. Apply to this state the following $(n\times n)$-matrices, which are also a subset of BST_n:

$$\begin{pmatrix} \frac12 & \frac12 &&&\\ \frac12 & \frac12 &&&\\ && 1 &&\\ &&& \ddots &\\ &&&& 1\end{pmatrix},\quad \begin{pmatrix} 1 &&&&&\\ & \frac12 & \frac12 &&&\\ & \frac12 & \frac12 &&&\\ &&& 1 &&\\ &&&& \ddots &\\ &&&&& 1\end{pmatrix},\dots,\begin{pmatrix} 1 &&&&\\ & \ddots &&&\\ && 1 &&\\ &&& \frac12 & \frac12\\ &&& \frac12 & \frac12\end{pmatrix},$$

$$\begin{pmatrix} \frac13 & \frac13 & \frac13 &&&\\ \frac13 & \frac13 & \frac13 &&&\\ \frac13 & \frac13 & \frac13 &&&\\ &&& 1 &&\\ &&&& \ddots &\\ &&&&& 1\end{pmatrix},\quad \begin{pmatrix} 1 &&&&&&\\ & \frac13 & \frac13 & \frac13 &&&\\ & \frac13 & \frac13 & \frac13 &&&\\ & \frac13 & \frac13 & \frac13 &&&\\ &&&& 1 &&\\ &&&&& \ddots &\\ &&&&&& 1\end{pmatrix},\dots,$$

$$\vdots$$

$$\begin{pmatrix} 1/n & 1/n & \cdots & 1/n\\ 1/n & 1/n & & \\ \vdots & & \ddots & \\ 1/n & 1/n & \cdots & 1/n\end{pmatrix}.$$

(Principle: The blocks are "growing" and "wandering" along the main diagonal.)

[1] For important hints concerning this procedure the author is indebted to Prof. Dr. J. Kerstan (Jena)

The new states obtained - something like "equalized partitions" - are called "child-points"[2] or more exactly child-points of first order. Child-points of first order child-points are child-points of second order, etc.

Those special child-points which are produced by matrices with $\frac{1}{2}$-blocks are called "daughter-points".

Then Kerstan could prove:

$$\varphi(\boldsymbol{p}) = \text{convex hull of } \boldsymbol{p} \text{ and its child-points of first order.}$$

3.3. $K^g(\boldsymbol{p})$ is built up

Connecting such "splinters" $\varphi(\boldsymbol{q})$ we construct the next set $K^g(\boldsymbol{p})$ (g for geometrical). The child-points determine how it happens:

$$\psi(1)(\boldsymbol{p}) := \varphi(\boldsymbol{p})$$

$$\psi(n+1)(\boldsymbol{p}) := \psi(n)(\boldsymbol{p}) \cup \left(\bigcup_{\{n\}} \varphi(\boldsymbol{q}) \right),$$

$\{n\}$ means that $\boldsymbol{q}$ runs over all child-points of nth order.

Then we define tentatively:

$$K^g(\boldsymbol{p}) := \lim_{i \to \infty} \psi^{(i)}(\boldsymbol{p})$$

But one can show that there is an integer M (M depends on the dimension, of course) with $\lim_{i \to \infty} \psi^{(i)} = \psi^{(M)}$. (The smallest of these numbers should be taken for M.) In other words: after a finite number of "procreations" the procedure becomes idempotent:

Lemma 3. *Let $\boldsymbol{p} \in \mathcal{S}_n$ and consider the sequence $\psi^{(i)}(\boldsymbol{p})$. Then $\exists M < \infty$, so that $\psi^{(M-1)} \subsetneq \psi^{(M)} = \psi^{(M+1)}$.*

Proof. At first we show that Lemma 1 is valid especially for child-points. (It is enough to consider those of first order.)
(i) For daughter-points it is clear, because $\boldsymbol{q} \in K(\boldsymbol{p})$ is valid by definition.
(ii) For the other ones we use Theorem 2‴. Let $\boldsymbol{q}$ be one of these, then by definition, $\boldsymbol{q} \in \varphi(\boldsymbol{p}) = K(\boldsymbol{p}) \cap E(\boldsymbol{p})$. Thus we have esp. $\boldsymbol{q} \in K(\boldsymbol{p})$.

This means that the assumptions of Lemma 1 are fulfilled and $K(\boldsymbol{q}) \subset K(\boldsymbol{p})$ holds for an arbitrary child-point $\boldsymbol{q}$ of $\boldsymbol{p}$.

A sequence $\{\boldsymbol{p}, \boldsymbol{q}', \boldsymbol{q}'', \dots\}$, appearing in the construction procedure, where the following point is a child-point of the former one is called a "line of ancestors". Now it could be seen that the creation of $K(\boldsymbol{p})$ by the sets $\psi^{(i)}$ is linked with a "tree" of such ancestor-lines and we study how it works in detail.

Starting with a single line we distinguish two cases: the line gets into a "new" cell, i.e. one in which no ancestor has been before or it gets into an old cell. In

[2] Because of the geometrical interpretation "state" and "point" are used as synonyms

the former case obviously a new contribution to $K(\boldsymbol{p})$ is added. In the latter case the following happens. Let $\hat{E}$ be this cell and $\tilde{\boldsymbol{q}}$ the ancestor of $\tilde{\tilde{\boldsymbol{q}}}$ that has already been in $\hat{E}$. We know that $K(\tilde{\tilde{\boldsymbol{q}}}) \subset K(\tilde{\boldsymbol{q}})$, thus $K(\tilde{\tilde{\boldsymbol{q}}}) \cap \hat{E}(\tilde{\tilde{\boldsymbol{q}}}) \subset K(\tilde{\boldsymbol{q}}) \cap \hat{E}(\tilde{\boldsymbol{q}})$, i.e. $\varphi(\tilde{\tilde{\boldsymbol{q}}}) \subset \varphi(\tilde{\boldsymbol{q}})$, and no new contribution appears. This is the mechanism of saturation.

In reality the construction of $K(\boldsymbol{p})$ results in a simultaneous creation of many ancestor-lines and we have to take into account their interactions. These will occur if two (or more, but it is enough to consider two) different lines get into a new cell. Let $\bar{\boldsymbol{p}}$ and $\bar{\boldsymbol{q}}$ be the last members of these two different lines just arriving at the new cell $\bar{E}$. Then we have neither $\varphi(\bar{\boldsymbol{q}}) \subset \varphi(\bar{\boldsymbol{p}})$ nor $\varphi(\bar{\boldsymbol{q}}) \supset \varphi(\bar{\boldsymbol{p}})$, but $\varphi(\bar{\boldsymbol{q}}) \cap \varphi(\bar{\boldsymbol{p}}) =: \Phi(\bar{\boldsymbol{p}}, \bar{\boldsymbol{q}}) \neq \varnothing$. Moreover, $\Phi(\bar{\boldsymbol{p}}, \bar{\boldsymbol{q}})$ is again of this splinter-shape, i.e. $\exists \bar{\boldsymbol{r}}$ with $\Phi(\bar{\boldsymbol{p}}, \bar{\boldsymbol{q}}) = \varphi(\bar{\boldsymbol{r}})$. By virtue of this point $\bar{\boldsymbol{r}}$ the two lines become comparable in the sense that one can switch into another line and can consider it as the proper one, more formally: let $\bar{\bar{\boldsymbol{p}}}$, $\bar{\bar{\boldsymbol{q}}}$ the respective child-points of $\bar{\boldsymbol{p}}$ and $\bar{\boldsymbol{q}}$, then $\bar{\bar{\boldsymbol{p}}} \in \varphi(\bar{\boldsymbol{p}})$, $\bar{\bar{\boldsymbol{q}}} \in \varphi(\bar{\boldsymbol{q}})$, $\bar{\boldsymbol{r}} \in \varphi(\bar{\boldsymbol{p}})$, $\bar{\boldsymbol{r}} \in \varphi(\bar{\boldsymbol{q}})$, $\bar{\bar{\boldsymbol{p}}} \in \varphi(\bar{\boldsymbol{r}})$, $\bar{\bar{\boldsymbol{q}}} \in \varphi(\bar{\boldsymbol{r}})$, and thus $\bar{\bar{\boldsymbol{q}}} \in \varphi(\bar{\boldsymbol{p}})$ and $\bar{\bar{\boldsymbol{p}}} \in \varphi(\bar{\boldsymbol{q}})$.

Because of the finite number of cells and the manner of procreation the last new cell is reached after a finite number of steps and in the next step the saturation as explained above happens. □

3.4. *The connection between $K^w(\boldsymbol{p})$ and $K^g(\boldsymbol{p})$*

Theorem 3. *Let $\boldsymbol{p} \in \mathcal{S}_n$. Then $K^w(\boldsymbol{p}) \equiv K^g(\boldsymbol{p})$.*

Proof. We take the definition of $K^g(\boldsymbol{p})$ and trace it back recurrently:

$$\begin{aligned} K^g(\boldsymbol{p}) &= \psi^{(M)}(\boldsymbol{p}) \\ &= \psi^{(M-1)}(\boldsymbol{p}) \cup \left(\bigcup_{\{M-1\}} \varphi(\boldsymbol{q}) \right) \\ &= \varphi(\boldsymbol{p}) \cup \left(\bigcup_{\{1\}} \varphi(\boldsymbol{q}) \right) \cup \left(\bigcup_{\{2\}} \varphi(\boldsymbol{q}) \right) \cup \cdots \cup \left(\bigcup_{\{M-1\}} \varphi(\boldsymbol{q}) \right). \end{aligned} \tag{$*$}$$

Considering $\boldsymbol{p}$ as order zero-child-point of $\boldsymbol{p}$ we can write compactly

$$K^g(\boldsymbol{p}) = \bigcup_{\{0 \cdots (M-1)\}} \varphi(\boldsymbol{q}) = \bigcup_{\{0 \cdots (M-1)\}} (K^w(\boldsymbol{q}) \cap E(\boldsymbol{q})). \tag{$**$}$$

Now we make use again of the fact that Lemma 1 is also valid for child-points. Thus $K^w(\boldsymbol{q}) \subseteq K^w(\boldsymbol{p})$ holds for an arbitrary child-point $\boldsymbol{q}$ of $\boldsymbol{p}$. (The equality sign ensures that zero order is also included.) Therefore we can replace all $K^w(\boldsymbol{q})$ by $K^w(\boldsymbol{p})$:

$$K^g(\boldsymbol{p}) = \bigcup_{\{0 \cdots (M-1)\}} (K^w(\boldsymbol{p}) \cap E(\boldsymbol{q})). \tag{$***$}$$

On the other hand we obviously have:

$$K^w(\boldsymbol{p}) = K^w(\boldsymbol{p}) \cap (\mathscr{S} \cup C\mathscr{S})$$
$$= K^w(\boldsymbol{p}) \cap ((\varepsilon^0 \cup \varepsilon^1 \cup \varepsilon^2 \cup \cdots \cup \varepsilon^{n!})).$$

The last bracket is to be understood as a "reservoir" of cells ε^j_n, so that there are enough of every kind. Now at first arbitrarily, we unite some ε^j_n into some $\tilde{E}^{\tilde{j}}$ and write

$$K^w(\boldsymbol{p}) = \bigcup_j (K^w(\boldsymbol{p}) \cap \tilde{E}^{\tilde{j}}).$$

We recall the definition of the $E(\boldsymbol{q})$: they just consist of unions of these ε^j_n, therefore we specify the $\tilde{E}^{\tilde{j}}$ just mentioned above to the $E(\boldsymbol{q})$ and arrive at

$$K^w(\boldsymbol{p}) = \bigcup_{\{0 \cdots (M-1)\}} (K^w(\boldsymbol{p}) \cap E(\boldsymbol{q})). \tag{****}$$

The comparison between (***) and (****) gives the statement. □

Therefore we set for future: $K^w(\boldsymbol{p}) \equiv K^g(\boldsymbol{p}) =: K(\boldsymbol{p})$.

In this proof there are also important hints as to the structure of $K(\boldsymbol{p})$. Some essential properties are provided by

Theorem 4. *Let $\boldsymbol{p} \in \mathscr{S}_n$. Then $K(\boldsymbol{p})$ has the following properties*:
(i) *$K(\boldsymbol{p})$ is a closed set.*
(ii) *$K(\boldsymbol{p})$ is starlike with respect to $\boldsymbol{e}$ (and therefore also connected).*
(iii) *$K(\boldsymbol{p})$ is in general not convex.*

Proof. At first we consider again (**) together with Theorem 2‴:

$$K(\boldsymbol{p}) = \bigcup_{\{0 \cdots (M-1)\}} \varphi(\boldsymbol{q}) = \bigcup_{\{0 \cdots (M-1)\}} (G(\boldsymbol{q}) \cap E(\boldsymbol{q})). \tag{**}$$

One can see that $K(\boldsymbol{p})$ is closed, because it is a finite union of a finite number of intersections of closed sets. Secondly one realizes starlikeness with respect to $\boldsymbol{e}$, since all $\varphi(\boldsymbol{q})$ are convex and contain $\boldsymbol{e}$.

Concerning non-convexity: if it is known that for $n = 3$ K($\boldsymbol{p}$) is not convex in general, then it is clear that also the higher-dimensional $K(\boldsymbol{p})$ must be so, since there are two-dimensional sections which are not convex. We give the proof for $n = 3$ explicitly by carrying out the construction given in Sect. 3.3. (See Figs. 4 and 5.) □

3.5. Remark

Not all details of the procedure for constructing $K(\boldsymbol{p})$ are known yet. Working with it one sees that there is some redundancy. One should expect that in the definition of the sequence $\psi^{(i)}(\boldsymbol{p})$ (see Sect. 3.3) $\boldsymbol{q}$ has only to run over the daughter-points (and not over all child-points). Moreover, one can prove that there are even daughter-points, which are unnecessary. Therefore we formulate

Problem 1. Find the minimal set of points necessary to construct $K(\boldsymbol{p})$!

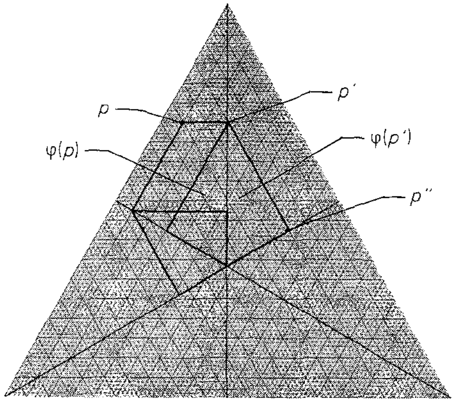

Fig. 4. Construction of $K(p)$ for $n=3$. The first two procreations: $\psi^{(1)}$ and $\psi^{(2)}$

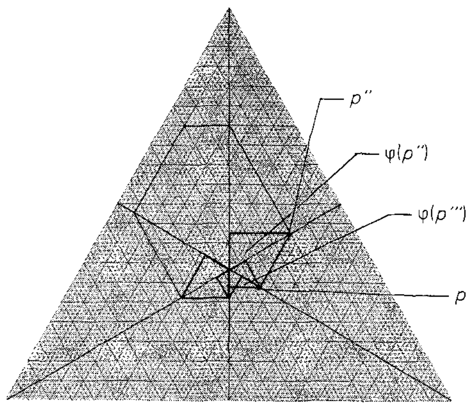

Fig. 5. Construction of $K(p)$ for $n=3$. Continuation and saturation: $\psi^{(3)}$ and $\psi^{(4)}$

Also open and interesting is

Problem 2. Find $M=M(n)$ (i.e. the minimal number of steps necessary to construct $k(p)$) for non-symmetric $p \in \mathcal{S}_n$! (For $n=3$, $M=4$ is true.)

4. The second result: another characterization of $K(p)$

Now we will present a further approach to our sets. Doing this we leave the special physical mechanism and replace it by well-defined mathematical objects.

A map $R^+ \ni t \to p(t) \in \mathcal{S}_n$ is called a process, for short $\{p(t)\}_{t\geq 0}$. If in addition we have $p(t'') \succ p(t')$, $\forall t'$, t'' with $0 \leq t' \leq t''$, then $\{p(t)\}_{t\geq 0}$ is called c-process (Lassner and Lassner [6]). (The "c" stands for "concave" and comes from an equivalent characterization of this partial order by concave functionals.) Being a subspace of R^{n-1} $\mathcal{S}_n$ is equipped with the induced topology and we can explain continuity: let $\{p(t)\}_{t\geq 0}$ be a c-process and in addition $\lim_{t\to t_0} p(t) = p(t_0)$, $\forall t$, $t_0 \geq 0$, then we call it a continuous c-process or cc-process.

Having this in mind, we define

$$K^{cc}(\boldsymbol{p}) := \{\boldsymbol{q} : \exists\ cc\text{-process } \{\boldsymbol{p}(t)\}_{t\geq 0}, \text{ so that } \boldsymbol{p}(0) = \boldsymbol{p} \text{ and}$$
$$\text{either } \boldsymbol{q} = \boldsymbol{p}(t') \text{ with } 0 < t' < \infty \text{ or } \boldsymbol{q} = \lim_{t\to\infty} \boldsymbol{p}(t)\}$$

This definition makes sense, because Alberti and Crell [1] proved that especially cc-processes always converge.

An easily demonstrable result is only mentioned:

Lemma 4. *Every w-process is a cc-process.*

Now we are going to prove

Lemma 5. *Let* $\boldsymbol{p} \in \mathcal{S}_n$ *and* $\{\boldsymbol{p}(t)\}_{t\geq 0}$ *a cc-process. Then follows* $\{\boldsymbol{p}(t)\}_{t\geq 0} \subset K(\boldsymbol{p})$.

Proof. This lemma is a consequence of both "c", namely (i) chaos-enhancing and (ii) continuous. (i) $\{\boldsymbol{p}(t)\}_{t\geq 0}$ should be a c-process, hence $\boldsymbol{p}(t) \in G(\boldsymbol{p}(0))$, $\forall t \geq 0$. (ii) From the existence of $E(\boldsymbol{p}(0))$ and the continuity of the process we get: $\exists T > 0$, so that $\forall t$ with $0 \leq t \leq T$ follows: $\boldsymbol{p}(t) \in E(\boldsymbol{p}(0))$. For these t (i) and (ii) provide: $\boldsymbol{p}(t) \in G(\boldsymbol{p}(0)) \cap E(\boldsymbol{p}(0))$. Because of Theorem 2‴ this means $\boldsymbol{p}(t) \in K(\boldsymbol{p}(0)) \cap E(\boldsymbol{p}(0))$ and esp. $\boldsymbol{p}(t) \in K(\boldsymbol{p}(0))$. Furthermore, Theorem 2‴ also guarantees that the argument just used for $t = 0$ can be applied to an arbitrary later t'. Thus there only remains to show that $\nexists T < \infty$ with the property $\boldsymbol{p}(T) \notin K(\boldsymbol{p}(t))$ for $t \leq T$, i.e. it is impossible for the trajectory to leave $K(\boldsymbol{p})$ at any time. For this we consider

$$Z := \{\infty > T > 0 : (1)\ \boldsymbol{p}(t) \in K(\boldsymbol{p}(0)), \forall t : 0 \leq t \leq T$$
$$(2)\ \exists \varepsilon > 0 : \boldsymbol{p}(T+\tau) \notin K(\boldsymbol{p}(0)), \forall \tau : 0 < \tau \leq \varepsilon\}$$

and $T_0 := \sup_{T\in Z} T$. Assume that $Z \neq \emptyset$, then there exist such T and because of the closedness of $K(\boldsymbol{p}(0)) = K(\boldsymbol{p})$, $T_0 \in Z$. The existence of $E(\boldsymbol{p}(0))$ guarantees: $\exists \delta > 0$ with $\boldsymbol{p}(T+\tau') \in K(\boldsymbol{p}(T_0))$, $\forall \tau'$: $0 \leq \tau' \leq \delta$ and therefore $\boldsymbol{p}(T_0+\tau') \in K(\boldsymbol{p}(0))$. But for an arbitrary $\tau' \in (0, \min\{\varepsilon, \delta\}]$ this provides a contradiction to condition (2), which T_0 as belonging to Z has to fulfill. Therefore Z must be empty, which means indeed that the trajectory always stays in $K(\boldsymbol{p})$. (The boundary is also allowed.) □

Now we show that $K(\boldsymbol{p})$ is just exhausted by the cc-processes:

Lemma 6. $\bigcup_{cc} \{\boldsymbol{p}(t)\}_{t\geq 0} = K(\boldsymbol{p})$

Proof. Considering Lemma 5 we have only to prove that every point of $K(\boldsymbol{p})$ is attainable by a cc-process. We mentioned above that every point of $K(\boldsymbol{p})$ can be reached by a finite number of two-body exchanges, and they are special cc-processes (compare Lemma 4). □

Therefore we can summarize:

Theorem 5. *Let* $\boldsymbol{p} \in \mathcal{S}_n$. *Then* $K^{w}(\boldsymbol{p}) \equiv K^{g}(\boldsymbol{p}) \equiv K^{cc}(\boldsymbol{p}) =: K(\boldsymbol{p})$.

These identities and the properties mentioned in Theorem 4 are the main results of this paper. More about this topic, especially about the $(n=3)$ - case can be found in [10].

5. Some consequences and further open questions

(1) We note that $G(\boldsymbol{p})$ is the set of states attainable from $\boldsymbol{p}$ by in general discontinuous c-processes. Therefore Theorem 2‴ means that $G(\boldsymbol{p})$ and $K(\boldsymbol{p})$ coincide *locally*. From a *global* aspect the additional requirement of continuity seriously restricts the set of attainable states, namely to $K(\boldsymbol{p})$.

Furthermore it makes sense to now introduce a sharper partial order than the "more mixed than" one in the following manner:

Assume that $\boldsymbol{p}, \boldsymbol{q} \in \mathcal{S}_n$. Then we define: $\boldsymbol{q} >> \boldsymbol{p}$ (in words: $\boldsymbol{q}$ is "more c-mixed than" $\boldsymbol{p}$), *iff* $\boldsymbol{q} \in K(\boldsymbol{p})$.

Obviously this relation is well adopted to consider continuous processes within $\mathcal{S}_n$, because permutations - as allowed in the "$>$"-concept and which destroy continuity - are forbidden here.

(2) The non-convexity of $K(\boldsymbol{p})$ is a surprise: from the mathematical point of view one knows and (thus) expects convexity arguments in this field. From a more physical point of view one is forced to overthink the intuitive picture of heat conduction attainability which we will illustrate in the following manner: consider two identical initial distributions $\boldsymbol{p}$, then apply to them two different w-processes T_1 and T_2, i.e. $T_1(\boldsymbol{p})=\tilde{\boldsymbol{p}}$ and $T_2(\boldsymbol{p})=\tilde{\tilde{\boldsymbol{p}}}$ with $\tilde{\boldsymbol{p}} \neq \tilde{\tilde{\boldsymbol{p}}}$, then mix $\tilde{\boldsymbol{p}}$ and $\tilde{\tilde{\boldsymbol{p}}}$ convexly, e.g. totally. You obtain two identical distributions $\boldsymbol{q} = \frac{1}{2}(\tilde{p}^1+\tilde{\tilde{p}}^1, \tilde{p}^2+\tilde{\tilde{p}}^2, \tilde{p}^3+\tilde{\tilde{p}}^3)$. Because of the non-convexity it could happen that $\boldsymbol{q}$ is *not* attainable from $\boldsymbol{p}$ although $\tilde{\boldsymbol{p}}$ and $\tilde{\tilde{\boldsymbol{p}}}$ of which $\boldsymbol{q}$ was mixed had been attainable. We give an example: $\boldsymbol{p}=(p^1, p^2, p^3)$,

$$\tilde{\boldsymbol{p}} = \left(\frac{1+p^3}{4}, \frac{3-p^3}{8}, \frac{3-p^3}{8}\right), \qquad \tilde{\tilde{\boldsymbol{p}}} = \left(\frac{3-p^1}{8}, \frac{3-p^1}{8}, \frac{1+p^1}{4}\right)$$

and

$$\boldsymbol{q} = \frac{1}{2}\left(\frac{5+2p^3-p^1}{8}, \frac{6-p^1-p^3}{8}, \frac{5+2p^1-p^3}{8}\right) \notin K(\boldsymbol{p}).$$

Problem 3. Is it possible to find an interesting "Gedankenexperiment" or even a technical application for this effect?

(3) The identity $K^w(\boldsymbol{p}) \equiv K^{cc}(\boldsymbol{p})$ can be taken as origin for considerations about the "replaceability" of cc-processes. Indeed this means that every cc-process can be replaced by a w-process in the following sense: assume a cc-trajectory from $\boldsymbol{p}$ to $\boldsymbol{q}$ with some arbitrarily chosen "check-points" on it. Then it is always possible to find a w-trajectory from $\boldsymbol{p}$ to $\boldsymbol{q}$ that contains all check-points. As an application of this we have: intuitively we might imagine any cc-process to be essentially a heat-conduction process. An example is

Lemma 7. *There is no cc-process that is also a cyclic one.*

A second application has already been tacitly used by us, namely that it is enough to consider only two-body exchanges: Heat exchange with more than two bodies involved is also a *cc*-process and therefore it can also be replaced by pure two-body exchanges. On the other hand one now clearly sees that the considerations on the special physical system presented in Sect. 1.2 were no more than a useful instrument to construct $K(\boldsymbol{p})$.

(4) As already mentioned, the class of *cc*-processes is a very large one. We complete the two examples from the introduction by adding the conditions for their coefficients here:

$$\text{(i)} \quad d/dt\, p^i = \sum_k L_{ik} p^k \text{ with } L_{ii} \leq 0, \forall i$$

$$L_{ik} \geq 0,\ i \neq k \text{ and}$$

$$\sum_i L_{ik} = 0, \forall k.$$

$$\text{(ii)} \quad d/dt\, p^i = \sum_{jkl} (A_{ijkl} p^k p^l - A_{klij} p^i p^j)$$

$$\text{with } A_{ijkl} \geq 0, \text{ and}$$

$$\sum_{ij} A_{ijkl} = \sum_{jk} A_{ijkl} = 1.$$

The Boltzmann-Carleman equation (ii) is remarkable because of its non-linearity (see [4]).

(5) Similar to one definition of $K(\boldsymbol{p})$ as the "two-body exchange future of $\boldsymbol{p}$" one can also consider the "two-body exchange past of $\boldsymbol{p}$". It turns out that this set is also non-convex.

We conclude with a remark concerning the matrix set K,

$$K := \left\{ T\colon T = \prod_i T_i,\ T_i \in T_{2(n)} \right\}.$$

This set is non-convex and non-closed.

Problem 4. Find out more about the structure of K!

Acknowledgement. For valuable discussions the author would like to thank Prof. Dr. A. Uhlmann and Dr. P. M. Alberti.

References

1. Alberti, P. M., Crell, B.: Boltzmannähnliche Gleichungen und H-Theoreme. Wiss. Z. Karl-Marx-Univ. Leipzig, Math. Naturwiss. R. **30**, 6, 539 (1981)
2. Alberti, P. M., Uhlmann, A.: Dissipative motion in state spaces. Leipzig: B. G. Teubner 1981
3. Alberti, P. M., Uhlmann, A.: Stochasticity and partial order. Berlin Dt.: Verl. d. Wissensch. 1981

4. Crell, B., Uhlmann, A.: An example of a non-linear evolution-equation showing "chaos-enhancement". LMP **3**, 463 (1979)
5. Hardy, G. H., Littlewood, J. E., Polya, G.: Inequalities. Cambridge: University Press 1934
6. Lassner, G., Lassner, G. A.: On the time evolution of physical systems. Dubna: Publ. JINR E2-7537, 1973
7. Muirhead, R. F.: Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters. Proc. Edinburgh Math. Soc. **21**, 144 (1903)
8. Ruch, E., Schönhofer, A.: Theorie der Chiralitätsfunktionen. Theoret. Chim. Acta (Berl.) **19**, 225 (1970)
9. Uhlmann, A.: Sätze über Dichtematrizen. Leipzig: Wiss. Z. d. Karl-Marx-Univ. **21**, 421 (1972)
10. Zylka, Ch.: Thesis (A) Leipzig 1982