Theoret. Chim. Acta (Berl.) 38, 167—183 (1975)

Commentationes

The Diagram Lattice as Structural Principle

A.
New Aspects for Representations and Group Algebra of the Symmetric Group

B.
Definition of Classification Character, Mixing Character, Statistical Order, Statistical Disorder; a General Principle for the Time Evolution of Irreversible Processes

Ernst Ruch

Institut für Quantenchemie der Freien Universität Berlin

Received November 30, 1974/December 23, 1974

It is the aim of this paper to demonstrate the significance of the diagram lattice. This lattice was defined in order to achieve structural insight into the phenomenon of chirality in chemistry. In this context, "Theorie der Chiralitätsfunktionen" [1] may serve as reference. In the introduction of the present paper a summary of the relevant theorems and definitions is given and a few examples of the diagram lattice are graphically illustrated. Parts A and B can be read independently and presuppose knowledge only of the introduction. Part A is of special interest for mathematicians, Part B and [1] for physicists and chemists.

In Part A theorems on the representations of the group $\mathfrak{S}_n$ and certain subgroups of it and on the structure of the group algebra are developed.

In Part B the concept "classification character" with the two complementary aspects of "identification" and "distinction" is derived. With the interpretation "mixing character" the partial order relation gains an interpretation through a mixing process, which can be expressed by a bistochastic matrix. This results in another equivalent definition of the diagram lattice. Interpreted as mixing character of a statistical distribution a diagram represents "statistical order" and "statistical disorder" by its row partition and column partition respectively. These concepts and the corresponding lattice structure lead to the hypothesis of growing mixing character as a criterion for the time evolution of isolated systems. The criterion of increasing entropy provides a much weaker condition. A discussion of the master equation leads to a proof of the principle of growing mixing character.

Key words: Symmetric group – Statistical order/disorder – Irreversible processes

The Diagram Lattice

In collaboration with A. Schönhofer [1], the author has introduced a "greater" relation for Young diagrams, in order to answer questions which appear in connection with the theory of chirality functions developed there. The partial ordering thus defined in the set of Young diagrams for any symmetric group $\mathfrak{S}_n$ applies to the set of irreducible representations of the group, as here is a one-to-one correspondence, as well as to that of the partitions of the integer n, i.e. to the various decompositions of n into a sum of integers. With the aid of this greater-relation, one can derive a number of theorems about the chirality of isomers and isomer mixtures, and find answers to questions which have proved relevant to chemistry and which could only be formulated in an appropriate way with reference to this set structure of diagrams. In our partial ordering, along with the "greater" relation, the lattice operations "union" and "intersection" are well-defined. For this reason we named the set of partitions of an integer the "partition lattice". For reasons which will become clear later, we wish to make use of the term "diagram lattice" when we concern ourselves with the set of diagrams.

Because of the usefulness of the structural concept, and because of the many-sidedness of its interpretation, it seemed probable to us that the diagram lattice might lead to many interesting insights in the field of pure mathematics and, in the form of applications, also in chemistry and physics. Since then, this impression has become stronger, and we deem it appropriate to assemble the aspects which we have found up to now, apart from those which are still in the realm of speculations.

One aspect which has encouraged us personally in our optimism in searching for new interpretations, even though it certainly is not a generally accepted scientific criterion is the beauty of this structure. This impression becomes evident through a graphical representation of examples, which also helps in the understanding of our definitions. We therefore consider it appropriate to present illustrations of the diagram lattice for several numbers. In this way, the concepts and theorems to be developed can be followed in terms of nontrivial examples.

We represent the partitions of a natural number in the usual way through diagrams, in which n boxes are arranged in rows such that the number of boxes in each row of a diagram characterizes the partition (see the illustrated examples). The arrangement of rows within a diagram in nondecreasing order of length from top to bottom is arbitrary from the standpoint of the partition, but nonetheless required in the sense of the representation through diagrams. We now formulate the "greater" relation by means of two different definitions, whose equivalence was proved in Ref. [1], and which will later be extended through other equivalent definitions.

Definition 1a. A diagram γ is called greater than a diagram γ', denoted by $\gamma \supset \gamma'$, if γ can be constructed from γ' by moving boxes exclusively upward, i.e. from shorter rows into longer or equal ones.

The ordering of the rows of a diagram from above according to non-increasing length implies a corresponding ordering of the columns, and *vice versa.* Thus, the "greater" relation can also be formulated in terms of the columns of the diagrams, as follows:

Definition 1b. A diagram γ is called greater than γ', denoted by $\gamma \supset \gamma'$, if γ' can be constructed from γ by moving boxes exclusively from right to left, i.e. from shorter columns into longer or equal ones.

The "greater" relation should also include equality, such that we have the conclusion

$$\gamma \subset \gamma', \quad \gamma \supset \gamma' \quad \Rightarrow \quad \gamma = \gamma' .$$

From the above definition, it easily follows that the transition from a diagram to a smaller neighboring diagram consists of moving one box into a shorter row.

Since the row- and column-lengths of a diagram each represent a partition of n – decomposition of n into a sum of integers – our ordering relation can be taken over immediately for partitions. In this connection, it only needs to be established whether the greater relation for diagrams should be taken over unaltered to the row- or to the column partition of the diagrams. We propose therefore

Definition 2. A partition is called greater than another, if the transition from the smaller to the greater can be made in steps, whereby partition numbers increase at the expense of others which are not larger.

From this follows

Definition 1c. A diagram γ is greater than a diagram γ', if and only if the row partition of γ is greater than that of γ', or in other words if the column partition of γ is smaller than that of γ'.

Definition 1c has been so formulated that for partitions one need not assume an ordering of partition numbers according to decreasing values. However, since the partition numbers are ordered in this way in diagrams, we define diagram partitions, and in particular row partitions and column partitions, as ones ordered with monotonically decreasing partition numbers. This ordering is not assumed for partitions without special designation.

From Definition 2, one can arrive at another formulation for the partial ordering of diagrams, in which the partial sums of the row- and column lengths play a role. We define v_i and μ_i respectively as the number of boxes in the i'th row and i'th column, and the partial sums

$$o_j = \sum_{i=1}^{j} v_i\,, \qquad u_j = \sum_{i=1}^{j} \mu_i\,.$$

In terms of the partial sums, one can formulate

Definition 1d.

$$\gamma \supset \gamma' \quad \Leftrightarrow \quad o_j \geqq o_j' \quad \Leftrightarrow \quad u_j \leqq u_j' \quad \text{for all} \quad j = 1, 2, \ldots n\,.$$

From this state of affairs one can prove the existence of the intersection and union, and one finds

$$\gamma \cap \gamma' = \gamma^D \quad \Leftrightarrow \quad o_j^D = \min(o_j, o_j')$$

$$\gamma \cup \gamma' = \gamma^V \quad \Leftrightarrow \quad u_j^V = \min(u_j, u_j')$$

These statements can be verified by examination of the lattice schemes for the numbers one through ten shown in Fig. 1. For proofs, the reader is referred to Ref. [1], where persuasive reasons for the formulation of the partition lattice are also to be found.

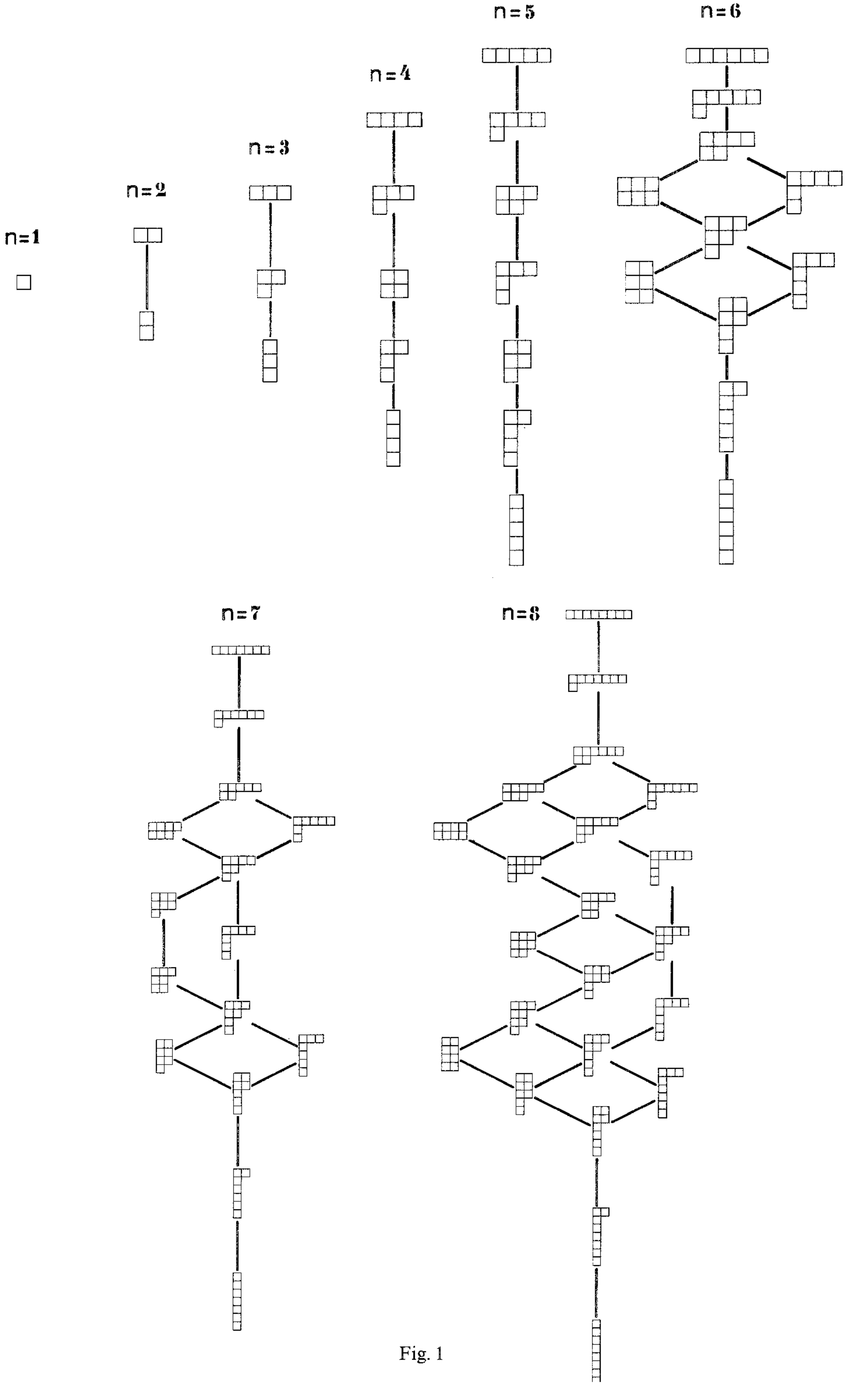

Fig. 1

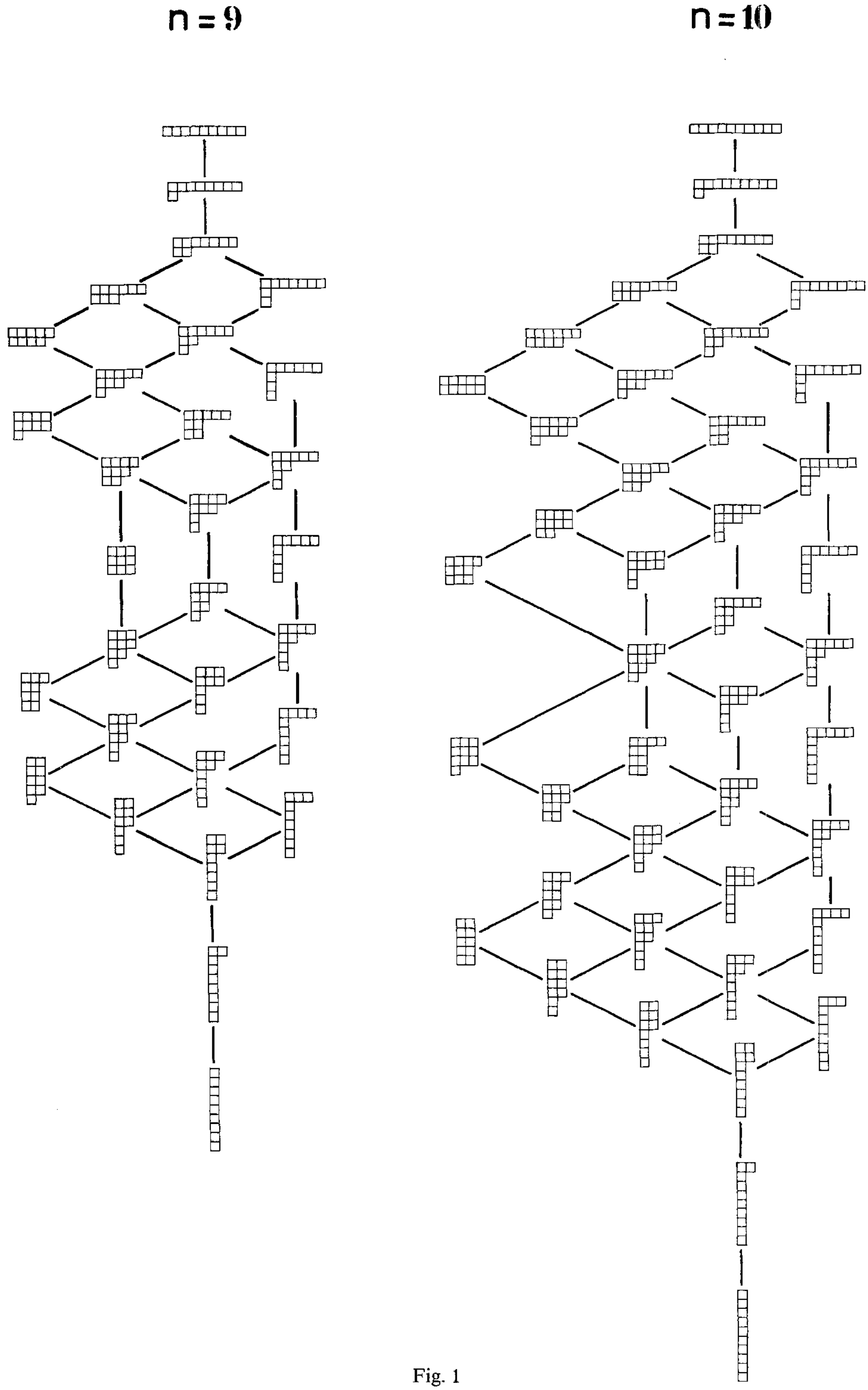

Fig. 1

Part A. The Diagram Lattice as Structure for Representations and Group Algebra of the Symmetric Group $\mathfrak{S}_n$

As shown by Young [2], there exists a one-one correspondence between the diagrams with n boxes and the irreducible representations of $\mathfrak{S}_n$. The correspondence is best understood through the irreducible left ideals, or through the simple two-sided ideals, of the group algebra. It is known that the irreducible left ideals as representation spaces give rise to irreducible representations, whereas the simple two-sided ideals give rise to representations whose irreducible parts are equivalent to one another. One forms tableaux, that is, one fills the boxes of a diagram with the numbers $1, \ldots, n$, and interprets the elements of $\mathfrak{S}_n$ as permutations of these numbers. A tableau-specific pair of subgroups $\mathfrak{A}_r$ and $\mathfrak{B}_r$ of $\mathfrak{S}_n$ is chosen, such that $\mathfrak{A}_r$ includes all permutations which do not remove numbers out of their rows, whereas $\mathfrak{B}_r$ includes all permutations of the numbers within the columns. Conjugate subgroup pairs belong to different tableaux of the same diagram, nonisomorphic subgroup pairs to different diagrams. We choose a particular tableau for each diagram, denote by $\mathfrak{a}_r$ and $\mathfrak{b}_r$ respectively the projection operators for the identity representation of $\mathfrak{A}_r$, and for the alternating (totally antisymmetric) representation of $\mathfrak{B}_r$:

$$\mathfrak{a}_r = \frac{1}{|\mathfrak{A}_r|} \sum_{s \in \mathfrak{A}_r} s; \quad \mathfrak{b}_r = \frac{1}{|\mathfrak{B}_r|} \sum_{s \in \mathfrak{B}_r} p(s)s$$

where $p(s) = +1$ for even and $p(s) = -1$ for odd permutations.

From these, one can construct the so-called Young operator y_r:

$$y_r = \mathfrak{a}_r \mathfrak{b}_r \quad \text{with the property} \quad y_r y_l \sim \delta_{rl} y_r .$$

It can be shown that y_r generates a minimal left ideal (and a minimal right ideal). In this way one finds a correspondence between irreducible representations Γ_r and diagrams γ_r:

$$\gamma_r \leftrightarrow \Gamma_r .$$

The two-sided ideal (y_r) generated by y_r is simple and is a sum of d_r equivalent left ideals. The representation defined by (y_r) decomposes into d_r equivalent irreducible representations Γ_r with dimension d_r. The ideal (y_r) contains a unit element, also called the projection operator for (y_r), of the form

$$p_r = \left(\frac{d_r}{n!}\right)^2 |\mathfrak{A}_r|\,|\mathfrak{B}_r| \sum_{s \in \mathfrak{S}_n} s y_r s^{-1} \quad \text{with the property} \quad p_r p_l = \delta_{rl} p_r .$$

The unit element of the group $\mathfrak{S}_n$, which is also the unit element of the group algebra, satisfies

$$e = \sum_{r=1}^{z} p_r$$

in which the sum runs over all irreducible representations (or diagrams).

It is desirable for a number of reasons – and not only for those reasons that we encountered in the theory of chirality functions – to obtain an overview of all those irreducible representations Γ_r of $\mathfrak{S}_n$ which contain the identity representation of $\mathfrak{A}_r$ or the alternating representation of $\mathfrak{B}_r$. As has been proved, though not explicitly emphasized, in (1, p. 279 ff.), the diagram lattice gives such an overview in the following way:

Theorem 1. The identity representation of the subgroup $\mathfrak{A}_r$ and the alternating representation of $\mathfrak{B}_r$ for an arbitrarily chosen tableau of the diagram γ_r, are contained in those irreducible representations Γ_l, and only in those, whose diagrams satisfy respectively $\gamma_l \supset \gamma_r$ and $\gamma_l \subset \gamma_r$. Neither representation of $\mathfrak{A}_r$ and $\mathfrak{B}_r$ is contained in representations whose diagram is not comparable with γ_r.

We remark that the diagram ordering introduced by Young in 1901, and since used in the mathematical literature, does not permit of such a statement.

There follows from theorem 1 a corresponding theorem about the group algebra, if one recalls the following facts:
The left ideals generated by $\mathfrak{a}_r$ and $\mathfrak{b}_r$ are representation spaces of $\mathfrak{S}_n$ with respective dimensions $n!/|\mathfrak{A}_r|$ and $n!/|\mathfrak{B}_r|$. The representations belonging to them, $\Gamma_{\mathfrak{a}_r}$ and $\Gamma_{\mathfrak{b}_r}$ decompose into irreducible parts Γ_l, and of these respectively $z_r(l)$ and $\bar{z}_r(l)$ are equivalent to one another. The left ideals are sums of minimal left ideals, of which respectively $z_r(l)$ and $\bar{z}_r(l)$ are operator-isomorphic to one another. $z_r(l)$ and $\bar{z}_r(l)$ simultaneously indicate how often respectively the identity representation of $\mathfrak{A}_r$ and the alternating representation of $\mathfrak{B}_r$ are contained in the irreducible representation Γ_l of $\mathfrak{S}_n$. In other terminology (cf. [1]), this means that the representations $\Gamma_{\mathfrak{a}_r}$ and $\Gamma_{\mathfrak{b}_r}$ are regularly induced from $\mathfrak{a}_r$ and $\mathfrak{b}_r$. An immediate consequence is

Theorem 2. The left ideals generated by $\mathfrak{a}_r$ and $\mathfrak{b}_r$ contain those minimal left ideals, and only those, whose diagrams γ_l satisfy respectively $\gamma_l \supset \gamma_r$ and $\gamma_l \subset \gamma_r$. The representations regularly induced from $\mathfrak{a}_r$ and $\mathfrak{b}_r$ have the decompositions

$$\Gamma_{\mathfrak{a}_r} = \sum_{\gamma_l \supset \gamma_r} z_r(l)\Gamma_l \qquad \sum_{\gamma_l \supset \gamma_r} z_r(l) d_l = \frac{n!}{|\mathfrak{A}_r|}$$
$$\text{where}$$
$$\Gamma_{\mathfrak{b}_r} = \sum_{\gamma_l \subset \gamma_r} \bar{z}_r(l)\Gamma_l \qquad \sum_{\gamma_l \subset \gamma_r} \bar{z}_r(l) d_l = \frac{n!}{|\mathfrak{B}_r|}.$$

The two-sided ideals $(\mathfrak{a}_r)$ and $(\mathfrak{b}_r)$ are direct sums of simple two-sided ideals of all diagrams γ_l satisfying respectively the conditions $\gamma_l \supset \gamma_r$ and $\gamma_l \subset \gamma_r$.

We note:
a) The general form of the elements of $(\mathfrak{a}_r)$ and $(\mathfrak{b}_r)$

$$(\mathfrak{a}_r):\quad \sum_{s,t \in \mathfrak{S}_n} a(s,t) s \mathfrak{a}_r t$$
$$(\mathfrak{b}_r):\quad \sum_{s,t \in \mathfrak{S}_n} b(s,t) s \mathfrak{b}_r t$$
with arbitrary complex coefficients $a(s, t)$ and $b(s, t)$

b) The unit elements of $(\mathfrak{a}_r)$ and $(\mathfrak{b}_r)$

$$(\mathfrak{a}_r):\quad e(\mathfrak{a}_r) = \sum_{\gamma_l \supset \gamma_r} \not{p}_l$$
$$(\mathfrak{b}_r):\quad e(\mathfrak{b}_r) = \sum_{\gamma_l \subset \gamma_r} \not{p}_l.$$

c) Special elements of $(\mathfrak{a}_r)$ and $(\mathfrak{b}_r)$

$$(\mathfrak{a}_r):\quad \bar{\mathfrak{a}}_r = \frac{1}{n!}\sum_{s\in\mathfrak{S}_n} s\mathfrak{a}_r s^{-1} = \sum_{\gamma_l\supset\gamma_r}\frac{z_r(l)}{d_l}\,\not{p}_l$$

$$(\mathfrak{b}_r):\quad \bar{\mathfrak{b}}_r = \frac{1}{n!}\sum_{s\in\mathfrak{S}_n} s\mathfrak{b}_r s^{-1} = \sum_{\gamma_l\subset\gamma_r}\frac{\bar{z}_r(l)}{d_l}\,\not{p}_l \qquad \text{with } z_r(r)=\bar{z}_r(r)=1\,.$$

From this it follows that

$$\bar{\mathfrak{a}}_r\bar{\mathfrak{b}}_s = \bar{\mathfrak{b}}_s\bar{\mathfrak{a}}_r \begin{cases} = \displaystyle\sum_{\gamma_r\subset\gamma_l\subset\gamma_s}\frac{z_r(l)\bar{z}_s(l)}{d_l^2}\,\not{p}_l \\ = 0 \qquad \text{if} \quad \gamma_r\not\subset\gamma_s \\ = \dfrac{1}{d_r^2}\,\not{p}_r \quad \text{if} \quad r=s. \end{cases}$$

From the above, one obtains without difficulty a correspondence between the set theoretical lattice of the two-sided ideals and the diagram lattice, as follows:

$$(\mathfrak{a}_r)\cap(\mathfrak{a}_s)=(\mathfrak{a}_t) \quad \text{where} \quad \gamma_r\cup\gamma_s=\gamma_t$$
$$(\mathfrak{b}_r)\cap(\mathfrak{b}_s)=(\mathfrak{b}_t) \quad \text{where} \quad \gamma_r\cap\gamma_s=\gamma_t\,,$$
$$(\mathfrak{a}_r)\cup(\mathfrak{a}_s)=(\mathfrak{a}_r,\mathfrak{a}_s)\subset(\mathfrak{a}_t) \quad \text{where} \quad \gamma_r\cap\gamma_s=\gamma_t$$
$$(\mathfrak{b}_r)\cup(\mathfrak{b}_s)=(\mathfrak{b}_r,\mathfrak{b}_s)\subset(\mathfrak{b}_t) \quad \text{where} \quad \gamma_r\cup\gamma_s=\gamma_t\,.$$

Here the symbols $\cap$, $\cup$, and $\subset$ denote the set theoretical intersection, union, and the statement "is a subset of...".

All this makes it clear that the diagram lattice contains genuinely meaningful relations regarding the structure of the group algebra and the induction of representations of $\mathfrak{S}_n$. Other algebraic consequences will be developed in a later paper (Part C).

Part B. Classification Character, Mixing Character and a Principle on Irreversible Processes

When we classify a set of n objects according to some principle, we obtain a subdivision into subsets without objects in common. Independently of the nature of the classification principle, we can represent the set structure of this subdivision by a diagram if, for example, we consider the rows as representing the numbers of equivalent objects. This, however, touches on only one side of the classification phenomenon, for the columns of the diagram give another aspect. The length of the leftmost column, μ_1, is the maximum number of nonequivalent objects; when μ_1 nonequivalent objects have been removed, μ_2 is the maximum number of nonequivalent objects in the remainder, etc. Both aspects, the "identification" and the "distinction" as characteristic features of a class subdivision represented by the row- and column-partitions, are equally valid designations of the classification structure in a set of n objects. Thus the lattice structure for the set of diagrams also has significance for the classification character, and the formulations for partitions

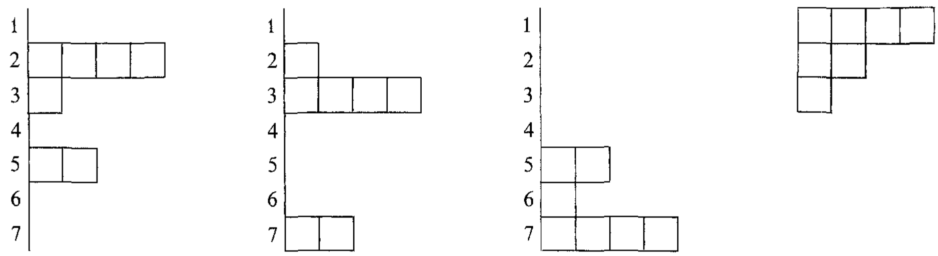

Fig. 2

according to definitions 1c and 1d correspond respectively to the identity and the distinction aspect. Summarizing, we make the following definition.

Definition 3. The classification character in a set of n objects is represented by a diagram. The complementary aspects "identification" and "distinction" correspond respectively to the row and column partitions. The following statements are equivalent: $\gamma \subset \gamma' \Leftrightarrow$ the distinction of the classification character γ is greater than that of $\gamma' \Leftrightarrow$ the identification of the classification character γ is less than that of γ'.

It is worthwhile to rederive this definition by reaching it in another way, not making use of our former lattice definition but analyzing the logical content of our notion. For this purpose we emphasize the viewpoint of distinction by translating classification character with mixing character – the mixing in the set of n objects is certainly a maximum if the number of distinguishable objects is maximal. The expression mixing character suggests, on the other hand, that increasing mixing character should be defined by a mixing process.

The mixing character thus denotes the composition of a set of partly equivalent objects, or the distribution of objects among different states. At this point, the question to be answered is whether, and to what extent, the concept of mixing character without arbitrary supplementary definitions (i.e., without use of phenomena other than that of mixing), can be used to compare sets of differing mixing character, such that a judgement of "more" or "less mixed" can be obtained, at least in comparable cases.

For such a concept, the nature of the objects or states is certainly not decisive, but rather the number of objects of the same kind or in the same state. Therefore, we call two sets "mixing-equivalent" if the partition of the classes is the same for both. It is sufficient to distinguish between just as many different kinds as the total number of objects in the set. With this restriction, we can denote all mixing-equivalent sets by a sequence of integers ν_i (including zero) in a column matrix, or graphically by means of diagram-like figures, in which rows with ν_i boxes are arranged along a vertical scale with indices $i = 1, 2, \ldots n$ denoting the different kinds of objects. By bringing the rows together into a diagram, we obtain the representation of the mixing character introduced above. Figure 2 gives an example of mixing-equivalent sets of seven objects, and of the representative diagram.

If the statement that a set $\mathfrak{M}$ is more mixed than another $\mathfrak{M}'$ is to have a meaning, the comparison must be restricted to pairs of sets such that $\mathfrak{M}$ is obtainable by mixing together sets of mixing character of $\mathfrak{M}'$. If further sets of mixing characters different from $\mathfrak{M}$ and $\mathfrak{M}'$ were necessary for the comparison, then the concept of a "more mixed" relation between two sets would have no meaning. Consequently, we must define $\mathfrak{M}$ as more mixed than $\mathfrak{M}'$ when $\mathfrak{M}$ can be obtained by mixing sets with the same mixing character as $\mathfrak{M}'$.

The naive mixing process consists in uniting whole multiples of mixing equivalent sets into a combined set containing a multiple of n objects. If the set finally obtained has a partition of objects which is an integral multiple of a partition of n, we may consider a corresponding set of n objects as equally mixed. This set may be regarded as being "more mixed" than any one of the sets which we used for mixing. As these have the same mixing character this is a relation between two mixing characters. The mixing procedure in this case may be expressed formally as a unification of mixing equivalent sets with non integral coefficients or, using an operator $\mathring{m} = \overset{\oplus}{\sum} c(s)s$ with the property $c(s) \geqq 0, \Sigma c(s) = 1$, as an operation on one of the mixing equivalent sets. The sum $\overset{\oplus}{\sum}$ should thereby be understood as the set-theoretical union of fractions of sets expressed by the coefficients, the permutations s effecting permutations of the object-types $1, 2, \ldots, n$. The corresponding group algebra element $\Sigma c(s)s$ acts on a column matrix with matrix elements representing the partition numbers of one of the mixing equivalent sets. Thereby the permutations s are represented by permutations matrices and the operator $\mathring{m}$ by a convex[1] linear combination of such permutation matrices. Since permutation matrices have a single element 1 in each row and in each column, with the rest of the elements being zero, the mixing operator is represented by a matrix M with the properties

$$\sum_i M_{ik} = \sum_k M_{ik} = 1 \quad \text{and} \quad M_{ik} \geqq 0 \quad \text{for all } i, k .$$

Such matrices are called bistochastic.

As is apparent from the above discussion, each convex linear combination of permutation matrices is a bistochastic matrix. The converse of this is almost trivial, since $n!$ permutation matrices of dimension n may be combined in many ways to give a specified bistochastic matrix. We thus arrive at the

Lemma. Mixing operators according to our definition are represented by bistochastic matrices, and conversely.

Through a theorem developed by G. H. Hardy, J. E. Littlewood and G. Polya [3], one finds a connection between the mixing operators and the partition lattice. The theorem refers to n-dimensional column matrices $\{x_i\}$ and $\{y_i\}$ of real numbers satisfying the condition $\sum_i x_i = \sum_i y_i$ and to partial sums of these numbers listed according to non-increasing values

$$x_{l_1} \geqq x_{l_2} \geqq x_{l_3} \geqq \cdots x_{l_n} \quad \text{and} \quad y_{l_1} \geqq y_{l_2} \geqq y_{l_3} \geqq \cdots y_{l_n} .$$

The partial sums shall be denoted by $X_k = \sum_{i=1}^{k} x_{l_i}$ and $Y_k = \sum_{i=1}^{k} y_{l_i}$. With this notation we formulate the

[1] A linear combination is called convex if $c(s) \geqq 0$ and $\Sigma\, c(s) = 1$.

Hardy, Littlewood and Polya Theorem. Of two n-dimensional column matrices $\{x_i\}$ and $\{y_i\}$ of positive real numbers satisfying the condition $\sum_i x_i = \sum_i y_i$ the first can be obtained from the second by multiplication with a bistochastic matrix if and only if relations of the form $X_k \leqq Y_k$ hold for all indices k, where the partial sums X_k and Y_k refer to the sequence of numbers according to decreasing values.

Since rows and columns in a diagram are ordered according to nondecreasing length, the mixing operator applied to a diagram partition transforms it into a smaller one. Moreover, every diagram partition is obtainable from every greater one by application of a mixing operator. It follows that the following definition is equivalent to 1a, 1b:

Definition 1e. $\gamma \subset \gamma' \Leftrightarrow$ the row partition of γ is obtainable from that of γ' by multiplication with a bistochastic matrix $\Leftrightarrow$ the column partition of γ' is obtainable from that of γ through multiplication with a bistochastic matrix.

Definition 3 follows inescapably if we replace mixing character with the more general concept classification character and substitute identification and distinction respectively for row and column partition.

Since bistochastic matrices are well-defined at least for denumerably infinite matrices and can evidently be defined for the continuum as well, definition 1d implies the natural generalization of the diagram lattice to the case of diagrams having an infinite number of boxes. By introducing a "box density", for example, one obtains an interpretation of the row partition as a monotone decreasing density distribution of boxes among a finite or denumerably infinite set of states, where the states are ordered according to decreasing occupation density.

Two concepts from statistical thermodynamics which are complementary in the same way as identification and distinction are those of "statistical order" and "statistical disorder". A distribution function in the usual sense represents the identification, since the number of objects in the same state is represented by the values of this function. A corresponding monotone decreasing function which we shall call "diagram distribution function" takes the place of a diagram with a finite number of boxes. It represents the statistical order as the row partition of a diagram does. The column partition or the inverse function, on the other hand, represents the sequence of decreasing maximum numbers of objects in different states. It indicates the statistical disorder in the distribution. Without going into topological questions which appear in the case of infinite sets, and which certainly possess only mathematical interest, we arrive at

Definition 4. The mixing character of a statistical distribution function is represented by a diagram function, the statistical order being designated by the row partition, and the statistical disorder by the column partition. Increasing mixing character is equivalent to increasing statistical disorder, and decreasing statistical order, and corresponds to the transition $\gamma \rightarrow \gamma' \subset \gamma$.

Every measurement of the mixing character according to some numerical scale corresponds to a homomorphic mapping of the diagram lattice onto the real numbers. In this mapping, incomparable diagrams are unavoidably mapped onto comparable values of the function used for measuring. Thus it is true that "greater mixing character leads to a greater value of the function", but not conversely. The entropy is such a homomorphic mapping of the partial ordering for the statistical disorder and the mixing character, as one can easily verify. The following proof makes use of definition 1a of the partial ordering.

What has to be shown is that

$$\gamma \Subset \gamma' \Rightarrow S(\gamma) > S(\gamma') \quad \text{where} \quad \gamma \Subset \gamma' \text{ means } \gamma \subset \gamma', \gamma \neq \gamma' .$$

For the proof, it suffices to show that the result is true for the case in which γ is obtained from γ' by the lowering of a single box. Thus, letting a single box be moved from the i'th row of γ' into the k'th leading to a new (unprimed) diagram γ, it follows that

$$(\nu_i + 1)^{\nu_i + 1} (\nu_k - 1)^{\nu_k - 1} > \nu_i^{\nu_i} \nu_k^{\nu_k}$$

and therefore

$$\begin{aligned} S(\gamma) - S(\gamma') &= - \Sigma\, \nu_s \ln \nu_s + \Sigma\, \nu_s' \ln \nu_s' \\ &= - \nu_i \ln \nu_i - \nu_k \ln \nu_k + (\nu_i + 1) \ln (\nu_i + 1) + (\nu_k - 1) \ln (\nu_k - 1) \\ &= - \ln \nu_i^{\nu_i} \nu_k^{\nu_k} + \ln (\nu_i + 1)^{\nu_i + 1} (\nu_k - 1)^{\nu_k - 1} > 0 . \end{aligned}$$

Since the greater relation for diagrams holds in the opposite sense for the mixing character, it follows that:

The mixing character of γ is greater than that of γ' $\Rightarrow$ $S(\gamma) \geqq S(\gamma')$.

At the same time that our Ref. [1] on chirality functions appeared, A. Uhlmann [4, 5] published a concept of "degree of mixture" of a density matrix which corresponds to the concept of mixing character developed here insofar as one thinks of mixing character in terms of the identification aspect.[2] The object of this work is a formulation of the Shannon entropy which does not depend on the spectral decomposition of density matrices. We make use of Uhlmann's result to apply our concepts to density matrices. According to the proofs given in Refs. [4, 5], a density matrix A is more mixed than a density matrix B if A is a convex linear combination of density matrices unitarily equivalent to B:

$$\begin{aligned} A \text{ more mixed than } B \quad &\Leftrightarrow \quad A = \sum_\mu c(\mu)\, U_\mu B\, U_\mu^{-1} \\ &\Leftrightarrow \sum_{i=1}^{j} \lambda_i^A \leqq \sum_{i=1}^{j} \lambda_i^B \quad \text{for all } j \end{aligned}$$

where $c(\mu)$ are real positive coefficients satisfying $\sum_\mu c(\mu) = 1$ and the U_μ are unitary matrices. The λ_i^A and λ_i^B are the eigenvalues of A and B in monotonically decreasing order. The proof of the equivalence of the above statement with the relation for the partial sums shows the equivalence of the two mixing concepts. The aspect of the column partition, distinction, or statistical disorder does not appear in Uhlmann's work, since diagrams play no role in his formulation of the problem. On the other hand, he recognizes the entropy as a homomorphic mapping of the partial ordering relation "more mixed than".

[2] The choice of the word "degree" seems to us unfortunate, since it implies in general the existence of a numerical scale. A scale on the other hand presupposes the lattice being modular, but this is not true for $n \geqq 7$ as one finds by inspection of the examples in Fig. 1.

A Theorem on the Time Development of Irreversible Processes in Closed Systems

The analysis of classification in a set of n objects at the beginning of this chapter led to the diagram as representative of the classification character. The diagram with its complementary decompositions in rows and columns leads inevitably to the equal roles of the two aspects "identification" and "distinction". With the interpretation "mixing character of a statistical distribution" for the diagram, "statistical order" and "statistical disorder" are the corresponding concepts for the row and column partitions. From an unprejudiced perception theoretical standpoint, it thus seems artificial not to consider both aspects in the same way in the formulation of numerical measures of the mixing character. Thus, if we decide to regard the Shannon entropy as an appropriate expression of statistical order and disorder, then we must grant the same role equally to a function which increases monotonically with mixing character like the Shannon entropy, but which is related to the column partition in the same way that the Shannon entropy is related to the row partition. Thus we compare two functions $S(\gamma)$ and $\bar{S}(\gamma)$ of the following form:

$$S(\gamma) = -\Sigma\, \nu_i \ln \nu_i\,; \qquad \bar{S}(\gamma) = \Sigma\, \mu_i \ln \mu_i$$

with the property

$$\gamma \subset \gamma' \Rightarrow \begin{cases} S(\gamma) > S(\gamma') \\ \bar{S}(\gamma) > \bar{S}(\gamma') \end{cases}$$

which can be proved for $\bar{S}(\gamma)$ in the same way as for the entropy, except that one considers columns instead of rows.

The following example (Fig. 3) shows that the order of the values of the two functions can be different for incomparable diagrams.

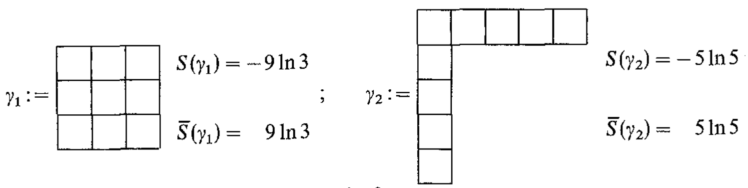

Fig. 3

It follows that

$$S(\gamma_1) < S(\gamma_2), \qquad \bar{S}(\gamma_1) > \bar{S}(\gamma_2)$$

and in this way it becomes clear that neither of the two functions alone, neither the Shannon entropy nor the function $\bar{S}(\gamma)$ provides a complete criterion for statistical order or disorder. $S(\gamma)$ and $\bar{S}(\gamma)$ are examples of what we shall call mixing homomorphic functions $f(\gamma)$, defined by the property

$$\gamma \subset \gamma' \quad \Rightarrow \quad f(\gamma) \geqq f(\gamma')\,.$$

Through examination of the examples of the diagram lattice given in Fig. 1, one can see that the three lowest diagrams are ordered in each case. If one limits oneself to these three lowest diagrams, therefore, each function $f(\gamma)$ with the property $\gamma \subset \gamma' \Rightarrow f(\gamma) > f(\gamma')$ also satisfies $\gamma \subset \gamma' \Leftrightarrow f(\gamma) > f(\gamma')$. Since the smallest diagram represents the uniform distribution, one can conclude that all functions of this kind have their maximum values at the equilibrium distribution for an isolated system and that, in the immediate neighborhood of equilibrium, all these functions represent the mixing character sufficiently by their monotonic dependence on it. Farther from equilibrium, this is no longer true. Hence, if one tries to make any mixing homomorphic function responsible for the temporal sequence of statistical distributions, it is important to realize that the sequence will depend on the choice of this function. From our standpoint, it is doubtful that the temporal sequence of distributions can be tied at all to any one function of this kind. As we have seen, the mixing character cannot be completely described by one function; the entropy is certainly a very vague criterion for situations far from equilibrium, while the increase in statistical disorder represents a generally acceptable standpoint. We feel that the interpretation given here of statistical order and disorder is quite persuasive, and it therefore seems reasonable to assume that the increase in statistical disorder, or of mixing character, is a general criterion for the temporal sequence of situations on the path to the equilibrium.

In the course of a thorough discussion of the manuscript of this work with my friends in the Chemistry Department, University of Minnesota, A. Mead made a suggestion which has helped us to formulate the above-mentioned hypothesis as a principle. For this, I owe Alden Mead my sincere thanks.

We consider the time development of a distribution of n particles among m states. The total number of particles is assumed constant, and the transitions are governed by constant transition probabilities w_{ik} in which the transition from k to i is denoted $(i \leftarrow k)$. The number n is assumed large enough that differentiation with respect to time is justified. This formulation is expressed through Eqs. (1)

$$\dot{n}_i(t) = \sum_{k=1}^{m} w_{ik} n_k(t) - \sum_{l=1}^{m} w_{li} n_i(t) \quad \text{where} \quad w_{ik} \geqq 0 \tag{1}$$

We rearrange (1) to

$$\dot{n}_i(t) = \sum_{k=1}^{m} W_{ik} n_k(t) \quad \text{where} \quad W_{ik} = w_{ik} - \delta_{ik} \sum_{l=1}^{m} w_{li} . \tag{2}$$

By defining vectors $\vec{n} = \{n_i\}$, we transform Eqs. (2) and their solution into the form

$$\dot{\vec{n}}(t) = W \vec{n}(t); \qquad \vec{n}(t) = e^{(t-t_0)W} \vec{n}(t_0) . \tag{3}$$

The matrix W satisfies

$$\sum_{i=1}^{m} W_{ik} = \sum_{i=1}^{m} w_{ik} - \sum_{l=1}^{m} w_{lk} = 0$$

that is, the sum of the elements in each column of W is zero.

We now specialize the conditions, such that the uniform distribution corresponds to the steady state as $t \to \infty$. The necessary condition for this, $mk = n$ (with integer k), is not critical for the condition on the system, because from $n_i(\infty) = n_k(\infty)$ for all i, k follows

$$\sum_{k=1}^{m} W_{ik} = 0$$

or: The sum of all the elements of each row of W is zero.

Thus, with the assumption of uniform distribution in the steady state, W, becomes a matrix in which the sum of the elements in each row and in each column is zero. We first convince ourselves that the product of two matrices whose rows and columns sum to zero also possesses this property, and then discuss the expansion

$$e^{(t-t_0)W} = E + (t-t_0)W + \tfrac{1}{2}(t-t_0)^2 W^2 + \cdots$$

where E is the unit matrix.

The expansion shows that the sum of matrix elements of $e^{(t-t_0)W}$ is one for each row and column, and it is evident that negative coefficients do not occur, because this would imply the possibility of negative $n_i(t)$. Therefore, we have the result that

$$e^{(t-t_0)W} \quad \text{is a bistochastic matrix,}$$

and consequently

$$\sum_{s=1}^{j} n_{l_s}(t) \leqq \sum_{s=1}^{j} n_{l_s}(t_0) \quad \text{for all} \quad j=1,\dots m \quad \text{if} \quad t \geqq t_0$$

where the increasing subindex s of l_s denotes monotonically decreasing values respectively of $n_{l_s}(t)$ and $n_{l_s}(t_0)$.
Thus according to definitions 1e and 4 we have the result.

The time development of the system proceeds in the direction of monotonically increasing mixing character, that is, monotonically increasing statistical disorder and monotonically decreasing statistical order.

This conclusion is considerably stronger than the known result that in such systems the (Shannon) entropy increases according to

$$-\Sigma n_i(t)\ln n_i(t) \geqq -\Sigma n_i(t_0)\ln n_i(t_0) \quad \text{for} \quad t \geqq t_0 .$$

Equation (1) is just a special formulation of the master equation, which has the general form

$$\dot{\varrho}(\omega,t) = \int [w(\omega,\omega')\varrho(\omega',t) - w(\omega',\omega)\varrho(\omega,t)]d\omega' \quad \text{with} \quad \begin{matrix} w(\omega,\omega') \geqq 0 \\ \int \varrho(\omega,t)d\omega = 1 . \end{matrix} \quad (4)$$

There $\varrho(\omega,t)$ denotes the statistical density of a Gibbs-ensemble, depending on the location ω in a space of variables, and on the time t. The transition probability from the point ω' to ω is expressed by $w(\omega,\omega')$. With the aid of the definition

$$W(\omega,\omega') = w(\omega,\omega') - \delta(\omega,\omega')\int w(\omega'',\omega)d\omega'' . \quad (5)$$

The masterequation (4) and its solution can be written in the form (6) and (7)

$$\dot{\varrho}(\omega,t) = \int W(\omega,\omega')\varrho(\omega',t)d\omega' \quad (6)$$

$$\varrho(\omega,t) = \int [e^{(t-t_0)W}](\omega,\omega')\varrho(\omega',t)d\omega' . \quad (7)$$

The known requirement on the system, to have an entropy increase with time is

$$\int w(\omega,\omega')d\omega = \int w(\omega',\omega)d\omega \quad \Rightarrow \quad S(t) \geqq S(t_0) \quad \text{for} \quad t > t_0 \quad (8)$$

where

$$S(t) = -k\int \varrho(\omega,t)\ln\varrho(\omega,t)d\omega \quad (k = \text{Boltzmann's constant}),$$

from this equation and the definition of $W(\omega, \omega')$ it follows that

$$\int W(\omega, \omega')d\omega' = \int W(\omega, \omega')d\omega = 0$$

and from the expansion

$$[e^{(t-t_0)W}](\omega, \omega') = \delta(\omega, \omega') + (t-t_0)\,W(\omega, \omega') + \tfrac{1}{2}(t-t_0)^2\,W^2(\omega, \omega') + \cdots$$

follows that

$$\int [e^{(t-t_0)W}](\omega, \omega')d\omega' = \int [e^{(t-t_0)W}](\omega, \omega')d\omega = 1$$

or

$$[e^{(t-t_0)W}](\omega, \omega') \qquad \textit{is a bistochastic function of } \omega \textit{ and } \omega'$$

or: the mixing character increases monotonically with time.

The master equation is applicable to a wide range of phenomena, probably wide enough to justify considering our hypothesis being a general principle, which we shall call the

Principle of Increasing Mixing Character

The time development of a statistical (Gibbs-) ensemble of isolated systems (microcanonical ensemble) proceeds in such a way that the mixing character increases monotonically.

Increase of mixing character is equivalent to increase of statistical disorder and decrease of statistical order, according to Definitions 3 and 4.

The most striking aspect of the principle of increasing mixing character, it seems to us, is that the evolution toward equilibrium is not determined by a quantity (such as the entropy, or any other single mixing-homomorphic function), but by a *quality*, the mixing character, which in turn is determined by the partial ordering defined by the diagram lattice.

Equal entropy for different mixing characters, if it is possible at all for a system with a finite set of particles, is very unlikely. Therefore, the time evolution of a system where all or nearly all mixing characters are passed whose entropy is not smaller than the initial one cannot be excluded by entropy arguments. The principle of growing mixing character on the other hand selects a comparatively small amount of different paths. Furthermore, with increasing n, incomparable mixing characters are to be found, which are separated from each other by increasing numbers of neighboring diagrams. This shows that the selection character of our principle compared with the principle of increasing entropy does not become negligible for large or infinite values of n.

It seems necessary to state that the discussion of equilibrium by means of arguments referring to entropy is unaffected by the principle of increasing mixing character. We have discussed this question for the case $m=n$ (p. 179). If $m>n$ the same argument applies as the mixing character is not dependent on the number of states but only on the maximum number of occupied states which is n again. If m is smaller than n we have to refer to diagrams with no more rows than m. For this a lemma is needed which is easily verified.

Lemma. The set of all diagrams with not more than m rows is a sublattice of the lattice with diagrams with $n>m$ boxes. This sublattice has a smallest diagram.

From the lemma follows that all mixing homomorphic functions have a maximum value at equilibrium.

It is desirable

1. to discuss our principle without reference to the master equation;

2. to find a set of mixing homomorphic functions according to the condition $\gamma \subset \gamma' \Rightarrow f(\gamma) > f(\gamma')$ which suffice to specify the increasing mixing character;

3. to give practical examples, which illustrate the relevance of the principle discussed above.

These problems will be attacked in subsequent papers.

The author wants to express his heartiest thanks to Professor A. Mead, who translated the German manuscript.

References

1. Ruch, E., Schönhofer, A.: Theoret. Chim. Acta (Berl.) **19**, 225 (1970)
2. Young, A.: Proc. London Math. Soc. (1) **33**, 97 (1901)
3. Hardy, G. H., Littlewood, J. E., Polya, G.: Inequalities 2nd Ed. Cambridge: University Press 1973
4. Uhlmann, A.: Rep. Math. Phys. **1**, 147 (1970)
5. Uhlmann, A.: Wiss. Z. Karl-Marx-Univ. Leipzig, Math.-Naturwiss. R. **20**, 633 (1971)

Prof. Dr. E. Ruch
Institut für Quantenchemie
Freie Universität Berlin
D-1000 Berlin 45
Holbeinstraße 48