

## THE WELL-REASONED CHOICE: AN INFORMATION-THEORETIC APPROACH TO BRANCHING RATIOS IN MOLECULAR RATE PROCESSES\*

R.D. LEVINE and R. KOSLOFF

*Department of Physical Chemistry, The Hebrew University,  
Jerusalem, Israel*

Received 11 June 1974

A collection of molecules has often a choice between competing (i.e., alternative) modes of evolution. It is customary to handle such problems through rate equations, thereby concluding that the branching ratio is determined by the relative rates of the competing processes. The grouping axiom of information theory is used to provide an alternative point of view, namely that the choice is based on the need to maximize the entropy, subject to known constraints. The thermodynamic interpretation and suggested applications are also considered.

### 1. Introduction

Chemical reactions, and rate processes in general, can often proceed along alternative (competing) paths [1]. The final outcomes will then differ. A reaction may, for example, yield either ground state or electronically excited products [2]. A large organic molecule promoted to an excited singlet state may fluoresce but can also undergo radiationless transitions to a triplet state [3]. An energy rich polyatomic molecule may isomerize or not. Many other examples readily come to mind. It is possible to characterize the preference (or lack thereof) for any particular route by a 'rate constant', such that if two routes are open, the fraction of products from the first route is just  $k_1/(k_1 + k_2)$ . While this provides a basis for quantitative measurements and the subsequent search for systematic trends, the interpretation (and ultimately, prediction) of such branching ratios from first principles is not always easy.

In this letter we outline an alternative approach to the interpretation of branching ratios. We take as the starting point the grouping axiom of information theory [4] and as our basic tenet that the branching ratio

is determined so as to maximize the entropy, subject to all the constraints in the problem. We shall moreover show that this approach can be justified using the concept of a thermodynamic weight [5]. In other words, we shall provide a thermodynamic interpretation of a branching ratio. In a sense to be made precise below, we shall show that the fractional yield (and hence the rate) for any particular path is proportional to the 'volume in phase space' accessible to it. Seemingly, we are making highly unreasonable claims. Yet, within our definitions, the claim is valid, with supporting experimental evidence.

A more limited version of the present approach has already been extensively documented and employed in the analysis of experimental results [6-8]. One could derive the present conclusions by an extension. We have chosen an alternative derivation since a modified approach often provides additional insight. In section 3 below we show how our previous results are a particular case of the present formulation.

### 2. The grouping axiom

To reduce the different physical problems to a common framework we choose the following schematic formulation [4]. The results of the experiment (say, many independent binary collisions) are divided into two

\* Acknowledgement is made to the donors of the Petroleum Research Fund, administered by the American Chemical Society for their support of this work.

(mutually exclusive) groups A and B. (Thus, A may be the set of collisions where reaction failed to occur and B the set of reactive collisions.) In general, we can, if we want to, resolve the events within each group into smaller 'bins'. (Thus, within group B, the group of reactive collisions, we can classify the outcomes according to the vibrational energy of the products, etc.) Let there be  $m$  different classifications ('bins') for the events in group A and  $n$  bins in group B. A particular outcome may thus be classified into any one of the  $n + m$  bins. We designate by  $P_i$  the fraction (i.e., the probability) of results that are classified into the  $i$ th bin,  $i = 1, 2, \dots, m + n$ . Taking the first  $m$  bins (i.e.,  $i = 1, 2, \dots, m$ ) as the group A, we have that the probability,  $p$ , of the group is given by

$$p = \sum_{i=1}^m P_i \quad (1)$$

In words, the probability of the group is the sum of the probabilities of the individual events in the group.  $p$  is the probability we want to determine.

Given that we have chosen an event from group A the probability that it is in the  $i$ th bin is

$$P_i^A = P_i/p, \quad i = 1, \dots, m, \quad (2)$$

which is just Bayes' theorem. Similar results apply to group B.

In general, the experiment will not lead to products all in the same bin. Rather, there will be a distribution of the outcomes. There will therefore be an uncertainty associated with the distribution [4]. We can then associate with the experiment an entropy  $S(P_1, \dots, P_{m+n})$  which is a function of the probabilities  $P_i$ . In the simplest case\* it is given by [6] ( $k$  is Boltzmann's constant)

$$S = -k \sum_{i=1}^{n+m} P_i \ln P_i \quad (3)$$

The discussion below does not require the explicit algorithm (3). All that is required is that  $S$  be a measure of the uncertainty (or lack of definitiveness) of the distribution of possible outcomes.

It is also possible to associate a measure  $S_A(P_1^A, \dots,$

$P_m^A)$  with the distribution within group A and similarly for group B. Explicitly

$$S_A = -k \sum_{i=1}^m P_i^A \ln P_i^A \quad (4)$$

$$\sum_{i=1}^m P_i^A = 1, \quad (5)$$

and similarly for group B.

The grouping axiom is the reasonable requirement that the uncertainty  $S$  associated with the experiment could be evaluated as the uncertainty associated with deciding between the two groups plus the (weighed) uncertainty of the two groups

$$S = S(p, 1-p) + pS_A + (1-p)S_B \quad (6)$$

Here  $S(p, 1-p)$  is the uncertainty which is eliminated when we specify either of the two groups (often termed 'the mixing entropy'). Explicitly

$$S(p, 1-p) = -k [p \ln p + (1-p) \ln(1-p)] \quad (7)$$

It is easy to verify that the explicit results (3), (4) and (7) do satisfy the grouping axiom.

For later reference we note that if we partition the outcomes into several groups  $\gamma$ ,  $\gamma = 1, \dots, r$  then

$$S = \sum_{\gamma} p_{\gamma} S_{\gamma} + S(p_1, \dots, p_r) \quad (8)$$

Here  $S_{\gamma}$  is the entropy for the distribution within the group  $\gamma$  and  $S(p_1, \dots, p_r)$  is the 'mixing' entropy

$$S(p_1, \dots, p_r) = -k \sum_{\gamma} p_{\gamma} \ln p_{\gamma} \quad (9)$$

The grouping axiom has a simple thermodynamic interpretation. Consider  $N$  outcomes, with  $N_j$  outcomes in the bin  $j$ ,  $j = 1, 2, \dots, m+n$ . Then the thermodynamic weight  $W_{AB}$  is [5]

$$W_{AB} = N! / \prod_{j=1}^{m+n} (N_j)! \quad (10)$$

In a similar fashion, if  $N_A$  outcomes belong to the group A, its thermodynamic weight is

$$W_A = N_A! / \prod_{i=1}^m (N_i)! \quad (11)$$

and similarly for group B. It is clear that

\* This applies when the index  $i$  refers to a single quantum state. If  $i$  refers to a group of states, a simple modification of (3) is required [6], as given in eq. (3i) below.

$$W_{AB} = W_A W_B W_M, \quad (12)$$

where  $W_M$  is the 'mixing' weight

$$W_M = N! / N_A! N_B! \quad (13)$$

If we now define

$$S = k \lim_{N \rightarrow \infty} (W_{AB} / N), \quad (14)$$

we recover (using Stirling's approximation) eq. (6), since, in the limit of a large  $N$ ,  $p = N_A / N$ .

### 3. The decision rule

The aim of the present note is to determine  $p$ , the fraction of outcomes which belong to classification A (e.g., electronic ground state reaction products). We propose that  $p$  is determined as that value which will maximize the entropy, subject to known constraints\*. As a simple example consider first the case when all the bins within the group are equally probable,  $P_i^A = 1/m$  and  $P_i^B = 1/n$ . Then  $S_A = k \ln m$  and similarly for  $S_B$ . Since

$$\partial S(p, 1-p) / \partial p = k \ln [(1-p)/p], \quad (15)$$

we obtain from (15) and (6) that the branching ratio is given by  $p/(1-p) = m/n$  or  $p = m/(m+n)$ . This conforms to the well tried rule of thumb of chemical kinetics that in the absence of any information the most reasonable assumption is that the branching ratio equals the ratio of the number of quantum states in the two groups (or, in classical language, the ratio of available volumes in phase space).

In the simple case we considered the group entropies were independent of  $p$ . In the general case they need not be. The general result obtained from (6) is

$$p/(1-p) = Z_A / Z_B \quad (16)$$

Here  $Z_A$ , defined by

\* In other words, if the predicted  $p$  agrees with the experimental results the constraints have been correctly identified. If not, there are additional constraints that need to be identified. Either way, one has gained insight.

$$k \ln Z_A = S_A + p \partial S_A / \partial p, \quad (17)$$

will be shown to be a partition function, of a particular type. Explicitly  $k \ln Z_A$  is the 'internal' contribution to the group entropy, in the sense of eq. (30) below.

The result (16) looks very much like thermodynamics. That is not surprising considering that it can be derived directly from (12) and (14). We recall that the thermodynamic weight  $W_{AB}$  is the 'number of complexions', i.e., the number of distinguishable results for given values of the distribution numbers  $N_j$ . When a very large number  $N$  of outcomes have been recorded, one particular set of distribution numbers  $N_j$  becomes much more probable. All we have done to obtain (16) is to determine  $p$  for that maximal  $W_{AB}$ . In other words (16) is the solution of  $\partial W_{AB} / \partial p = 0$  in the limit  $N \rightarrow \infty$ . This procedure is not restricted to equilibrium situations nor is it applicable only to molecules†.

We turn next to the explicit role of constraints.

### 4. Constraints

To determine the branching ratio one varies the probabilities of the individual outcomes so as to achieve an extremum of  $S$ . The  $n+m$  different probabilities  $P_i$  cannot be varied as independent variables due to the ever-present normalization condition. There are thus at most  $n+m-1$  independent variations. Rather than vary the  $P_i$ 's we can vary  $p$  and the probabilities within the groups ( $P_i^A$  and  $P_i^B$ ). Since the probabilities within each group are normalized, eq. (5), there are  $(m-1) + (n-1)$  independent variations within the groups, plus the variation of  $p$ ,  $n+m-1$  in all. In other words, when we vary  $p$ , we need to impose the normalization condition on each group separately.

The variations in the probabilities may be further constrained by the need to conserve not just the normalization but also the expectation value of other, non-trivial quantities (e.g., the energy). We can distinguish two possibilities. The first, and simpler case, is when the average value of some attribute is specified separately for the two groups. Thus if  $g(i)$  is the value of

† Consider throwing  $N$  dice together. Surely as  $N \rightarrow \infty$ , the branching ratio for even versus odd faces settles down (to 1/2 in the absence of information suggesting loaded dice or otherwise).

property  $g$  for the state  $i$ , its average value for the group A is

$$\langle g \rangle_A = \sum_i g(i) P_i^A. \quad (18)$$

It is important to note that the value of  $\langle g \rangle_A$  is unchanged when we vary  $p$  and similarly for  $\langle g \rangle_B$ . One can readily show that if normalization and a given value of  $\langle g \rangle_A$  are imposed as constraints  $S_A$  will be a function of  $\langle g \rangle_A$  [5, 9] and hence will be independent of  $p$ . A similar conclusion holds when we constrain several averages (say  $\langle g_r \rangle_A$ ,  $r = 1, 2, \dots$ ) over group A and similarly for group B.

When the constraints apply separately to the two groups,  $\partial S_X / \partial p = 0$ ,  $X = A, B$ , and (16) yields for the branching ratio

$$p/(1-p) = \exp[(S_A - S_B)/k]. \quad (19)$$

If only intragroup constraints are invoked, each group can be handled separately and the usual expressions for entropy apply. Explicitly [5, 9]

$$S_A = k \sum_r \lambda_r \langle g_r \rangle_A + k \ln Q_A, \quad (20)$$

$$Q_A = \sum_i \exp \left[ - \sum_r \lambda_r g_r(i) \right] \quad (21)$$

and the  $\lambda_r$ 's are determined by

$$\langle g_r \rangle_A = - \partial \ln Q_A / \partial \lambda_r. \quad (22)$$

Similar results hold for group B.

When there are only intragroup constraints [of the type (18)] we have thus obtained an explicit expression for the branching ratio, (19), in terms of the known averages  $\langle g_r \rangle_A$  and  $\langle g_r \rangle_B$ . The algorithm is clear. Consider the partition function  $Q_A$ , eq. (21), as a function of the  $\lambda_r$ 's. Use the known values of  $\langle g_r \rangle_A$ ,  $r = 1, 2, \dots$  to determine the magnitude of the  $\lambda_r$ 's, eq. (22). Use these values in eq. (20) to determine  $S_A$ . Repeat for group B.

The second type of constraint affects both groups. Here we only constrain the average value of some attribute  $E$  over the entire set of outcomes

$$\langle E \rangle = \sum_{i=1}^{m+n} E_i P_i = p \langle E \rangle_A + (1-p) \langle E \rangle_B. \quad (23)$$

The most typical example of such a constraint is a pro-

cess which is carried out isothermally. Here only the average value of the energy is prescribed, but not the fraction of energy within each group. Thus, in the presence of an intergroup constraint, we seek the maximum of  $S$  subject to  $d\langle E \rangle = 0$ . However, we do not constrain  $d\langle E \rangle_A$  nor  $d\langle E \rangle_B$ . To search for the constrained maximum we introduce a Lagrange multiplier  $\beta$  and require that  $dS - k\beta d\langle E \rangle = 0$ . Using eqs. (6) and (23), the stationary value of  $p$  is determined from

$$\partial S(p, 1-p) / \partial p + S_A - S_B - k\beta(\langle E \rangle_A - \langle E \rangle_B) = 0. \quad (24)$$

To complete the argument we differentiate (23) with respect to  $p$  holding  $\langle E \rangle$  constant. Then

$$\langle E \rangle_A - \langle E \rangle_B = -p \partial \langle E \rangle_A / \partial p + (1-p) \partial \langle E \rangle_B / \partial (1-p). \quad (25)$$

Finally, when  $\langle E \rangle$  is imposed as a conserved value in addition to the intragroup variables, (20) is modified to

$$S_A = k\beta \langle E \rangle_A + k \sum_r \lambda_r \langle g_r \rangle_A + k \ln Q_A. \quad (26)$$

Thus, using (25) and (26)

$$k\beta(\langle E \rangle_A - \langle E \rangle_B) = p \partial S_A / \partial p - (1-p) \partial S_B / \partial (1-p). \quad (27)$$

Substituting (27) in (24) completes the expression for the branching ratio

$$p/(1-p) = Z_A / Z_B. \quad (28)$$

Our derivation has provided an explicit algorithm for the  $Z$ 's,

$$k \ln Z_A = S_A - k\beta \langle E \rangle_A. \quad (29)$$

If there are no intragroup constraints,  $Z_A$  is simply the partition function for group A. Otherwise, using (26)

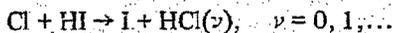
$$k \ln Z_A = k \ln Q_A + k \sum_r \lambda_r \langle g_r \rangle_A, \quad (30)$$

and this result is valid in general, irrespective of the number of distinct intergroup constraints.

The explicit algorithm for  $Z_A$  and hence for the branching ratio is the central result of this note. One can either interpret  $Z_A$  as a partition function [cf. eq. (29)] or, from eq. (30) as that contribution to the entropy which is 'internal' to the group.

## 5. Past and future applications

The analysis of alternative modes of energy disposal [6] can be readily rephrased as an illustration of the present approach. We consider the different product vibrational states in a reaction, e.g.,



as representing different reaction paths so that  $P(\nu)$  is the fraction of products where HCl is in the  $\nu$ th vibrational state. In the absence of information to the contrary we assume that all quantum states of the products in the  $\nu$ th group are equally probable. In other words, we assume no intragroup constraints so that the entropy of the group  $\nu$  is  $S_\nu = k \ln g_\nu$ , where  $g_\nu$  is the number of quantum states in the group. Then (8) can be written as

$$S = -k \sum_{\nu} P(\nu) \ln [P(\nu)/g_{\nu}], \quad (31)$$

and the weights  $P(\nu)$  are to be determined from intergroup constraints. This indeed is the procedure previously recommended and employed [10, 6-8]. The only change is a change in interpretation. In the present approach we are quite explicit in recognising that the weights of the groups reflect the outcome of competing reaction paths. There is therefore no inherent

reason why a similar analysis cannot be applied to more complicated situations involving both inter- and intragroup constraints. Indeed an application to radiationless transitions (where radiative and non-radiative decay modes compete) has been preliminarily reported [11] and other applications are in progress.

## References

- [1] R.D. Levine and R.B. Bernstein, *Molecular reaction dynamics* (Clarendon Press, Oxford, 1974).
- [2] M. Menzinger and D.J. Wren, *Chem. Phys. Letters* 18 (1973) 431.
- [3] J. Jortner, S.A. Rice and R.M. Hochstrasser, *Advan. Photochem.* 7 (1969) 149.
- [4] R. Ash, *Information theory* (Interscience, New York, 1965) p. 8.
- [5] R.D. Levine and R.B. Bernstein, *Chem. Phys. Letters* 22 (1973) 217.
- [6] A. Ben-Shaul, R.D. Levine and R.B. Bernstein, *J. Chem. Phys.* 57 (1972) 5427.
- [7] R.D. Levine, B.R. Johnson and R.B. Bernstein, *Chem. Phys. Letters* 19 (1973) 1.
- [8] A. Ben-Shaul, *Chem. Phys.* 1 (1973) 244.
- [9] E.T. Jaynes, *Phys. Rev.* 106 (1957) 620.
- [10] G.L. Hofacker and R.D. Levine, *Chem. Phys. Letters* 15 (1972) 165.
- [11] R. Kosloff, R.D. Levine and O. Kafri, *Proceedings Radiationless Transitions Conference, Schliersee* (1974).