Thermochimica Acta, 187 (1991) 159–166
Elsevier Science Publishers B.V., Amsterdam

The Theorem of Common Kinetic Runs – a concept against an unorganized data overflow in reaction kinetics

E. Koch

Max–Planck–Institut für Strahlenchemie
Stiftstr. 34–36, 4330 Mülheim an der Ruhr, Germany

Abstract

Based on the well–proven system of kinetic coordinates, derived from shape index and halfwidth of a derivative non–isothermal signal, on the activation data of the internal one–step reference reaction and on N intervals of the starting concentration of a reactant, a 3 x N matrix can be defined which fixes the kinetics occurring in a homogeneous system and is available from experimental or theoretical series. This readily available matrix, called Mechanistic Concentration Code (= MCC), was recently proven to be an ideal vehicle for data abstraction because it is nearly independent of activation data and methodical coefficients and indicates the rate–determining source processes and their molecularities. As was confirmed by DTA– and rate curves of numerous complex reactions, it enables one to get a distribution list for all two–step processes involved, revealing smallest differences in the kinetics.

1. INTRODUCTION

Because of the paramount evolution of data processing, the natural scientist is exposed to the danger of getting drowned by a self–generated data flow if an adequate filter for the information is missing /1–3/. Therefore, especially in Thermal Analysis /4/, where temperature is an additional parameter, a data abstraction directly after a time–resolved experiment becomes increasingly important. Its principle should be as transparent as possible for the user, in order to prevent that he will be exposed to a "black box" apparatus.

The primary aim of kinetic measurements is to find out the best model ruling the temporal behaviour of the system investigated. In complex homogeneous kinetics in solution, the building stones of such a model are uni– or bimolecular reaction steps. Thus the question arises whether there are typical parameters or parameter sets which characterize the model, independently of rate constants or activation parameters.

2. MECHANISTIC CONCENTRATION CODES

Indeed it was found that such a set really can be defined and obtained from kinetic experiments. Based on a series of runs, performed in N intervals of the starting concentration of the experiments, the kinetic information, representing the network of the rate–determining steps and their molecularities, is condensed most effectively by a 3 x N matrix (= MCC) for each of p–1 components of a system involving p reacting species. The three lines of a MCC are formed from elements denoting behaviour types which refer to three model–specific response parameters being approximately invariant for elementary processes /5,6/:

0040-6031/91/$03.50 © 1991 Elsevier Science Publishers B.V., All rights reserved

1. Shape index (asymmetry of the derivative signal)
2. Initial Reaction type index (relative signal height with respect to the start reaction as reference /7/; cf. /8,9/)
3. Overall Reaction type index (the same for the fitted overall reaction).

The simplest elements of such a MCC are '+' for increasing, '–' for decreasing, and 'C' for constant reponse parameter when the starting concentration is increased. The 'C'–elements correspond to the (often apparent) molecularity of the correspondent rate–determining step in the interval considered. Therefore, they were classified further due to the mathematical order (this is the molecularity, if n reactant concentrations are equal in a n–molecular process): L(ow),1, M(edium),2, H(igh).
Such a matrix meets with the requirements of an ideal kinetics data abstraction because rate–determining steps and their order are fundamental in every discussion on kinetics, similarly as the concepts of the initial rate and the overall reaction.
Let us consider the simplest composite models, the group of two–reaction models (= TRM's). If the three strings of a MCC are standardized to unit conversion and to fixed limits of the starting concentration (c_0 = 0.00005, c_∞ = 100 mole/dm^3 are generally sufficient), it can generally be reached by fit of the second A–factor that – independent of the remaining three activation data – the behaviour of the system converges at both ends towards a constant element, due to the corresponding source reaction. Based on the concept of their creation, the strings are independent of the activation parameters. If one of the four activation parameters is changed, only a common shift of the three strings may occur towards the strings of a permanent, experimentally gained MCC. This shift is caused by the so modified reference point of the concentration scale, depending on the specific times /10,16/ of both reactions.

3. PROOF OF A "THEOREM OF KINETIC RUNS"?

Usually, there is no distinct influence of the activation data; but there are exceptions. For example, for the model P22 (two concurrent second–order steps) elements inside the string may degenerate since this model might scarcely be identified from a concentration–based study. In other (unprobable) cases, as for extremely different activation energies, apart from such degeneration effects even the appearance of new elements may be stated. In certain consecutive models, the final stationarities cannot be reached for technical reasons. Therefore, numerous tests were planned, using some special differences of the activation energies (E_2–E_1), namely between –10 and +20 kcal/mole (*inverse* (<0), *equal*, *small*, *wide*, *normal*, and *equal* cases). For standardization, E_1 = 15 kcal/mole and log A_1 = 12 were mostly assumed (*normal*); taking $(\log A)_1$ = 9 (*slow*) or $(\log A)_2$ = 12 (*fast*) instead had no influence. Other changes of the strings were in general negligible.

A much more critical problem seemed to be the influence of the *method of measurement*. In a kinetic study, the method can often be characterized by proportionality factors for the particular rates of the steps /5/; this influence has been studied in a large–scale computer project /10/. Since pre–studies had revealed that also this influence is relatively small, these *relative signal parameters* λ_2/λ_1 were restricted to the following values: 0, 0.3, 1.0, 3.0, ∞ (and –0.3 in special models). In this sequence, the role of the second reaction in the direct observation is increased from zero whilst that of the first reaction (reference) is reduced, until for $\lambda_2/\lambda_1 = \infty$ only reaction 2 remains for the observation (however, the kinetic involvement of both reactions – controlled by their specific times – is not changed).

The range of possible parameters is restricted by *thermodynamical relationships*. Thus, for opposing reactions (TRM's G11, G12, GAB1 etc.) λ_2/λ_1 = 1 must be

assumed (*reversibility*). Further, for a definite type of transition complex, the A–factor is fixed /11/, and activation energy and reaction enthalpy may satisfy a definite relationship /12/. On the other hand, in the case of an unknown model and absence of reversible steps, the restrictions may be rather poor; the only conclusion is that strongly positive enthalpies can be omitted. Finally, in small models such as irreversible TRM's there are no branchings or circuits between two species, so that also stoichiometric equations between the signal parameters are missing.

For all these reasons, a certain variation of the *signal parameter* could not be dispensed with. Thus, λ_2/λ_1 = 0.3, 1 and 3 were tested for all 25 TRM's, but –0.3, zero and infinite only in somewhat doubtful cases. Despite these restrictions, the considerable problem remained to study ca. 250 series based on at least 15 different starting concentrations, according to more than 3000 theoretical plots. These were created in a half–year period, after the development of a highly automatic computer program /10/.

The results of these fundamental studies, generation and analysis of the MCC's of all two–step models revealed a common architecture, characterized by three types of elements:

1. *Limiting elements.* The first and last elements in a string, which correspond to very low and to very high initial concentration, usually indicate constancy. In most cases, the unimolecular or bimolecular source reactions (elements '1' and '2') can be directly correlated; in other cases, the corresponding orders are fractional.

2. *Main elements.* Between the limiting elements, there are 2 – 6 elements of the '–' or '+' type. Their sequence is typical for the prevailing model and does not essentially depend on the signal– or activation–parameters.

3. *Rest elements.* Apart from the elements mentioned, there may be elements of all types which for a definite model may differ for the three "standard" signal parameters scheduled before (0.3, 1 and 3). In many cases, they appear only for either λ_2/λ_1 = 1 or (inversely) for λ_2/λ_1 = not 1 (as a kind of "stoichiometric resonance effect"). Since there is no third source reaction, 'C'–elements in this group belong rather to a slightly bent, quasiconstant type around a maximum or minimum.

For the exact validity of the above "Theorem", there should be no *rest elements*, as generally for the *G–types*. In reality, the contributions of the *rest elements*, evaluated statistically over all 250 series and 50 models /10/, are distinct, but relatively small; approximately

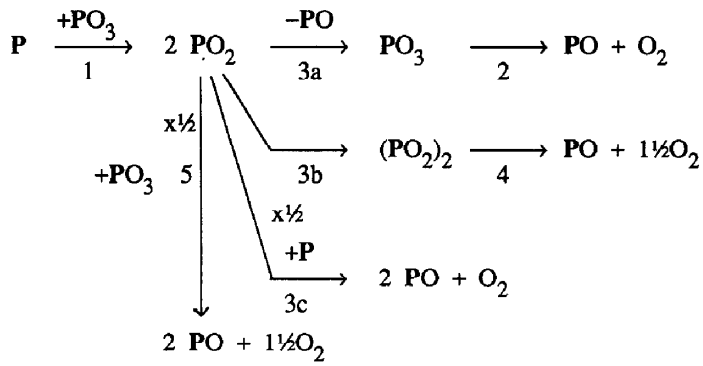
10% of the elements (3 x N for each model) for parallel reactions (= P),
16% for consecutive (= F), and
20% for independent reactions (= U).

These percentages are reduced further, if in the case of several nearly isolated peaks, separate MCC's are provided for the peak groups. This situation often holds for F– or U–types or for those rather extraordinary P–types where the signal parameters show different signs.

4. APPLICATIONS OF PATTERN RECOGNITION

The approximate validity of the "Theorem" admits for every experimental or theoretical series of kinetic plots to create a distribution list of all TRM's at all, assorted with respect to their probabilities. The rating procedure has been described in

detail /6/; a review of all needed expressions for an evaluation is in preparation for this journal. More than 100 DTA series of various reactions in solution have been studied, confirming that the procedure is sensitive to smallest changes of the kinetics.



This is demonstrated looking at a complicated reaction of simple reactants where at least 6 steps have to be considered. The ozone adduct PO_3 of triphenyl phosphite (= **P**) reacts in methylene chloride /13,14/ with additional **P**, causing up to four DTA peaks. The first (= α) is caused by the scavenging reaction of PO_3 with exceeding **P**; consequently, this peak is shifted to lower temperature if $[P]_0$ is increased, whereas the third peak (= γ), responsible for the unimolecular decomposition of PO_3, holds its temperature position. The β– and δ–peak may be understood similarly via the formation of dimer–type components; see scheme and Table 1 where the corresponding TRM's are listed; reference component: **P** (varied); reference reaction: No.1.

Table 1
Correlations of two–step models to the steps in the scheme (POZ6)

TRM	Reaction(s)	evidence factor
AB	1	23.8%
PAB1	1 + 2	36.7% *)
PBA1	1 + 2, but PO_3 as reference component	46.8%
P21	1(fast) + 3b + 3c (strong **P**–excess)	42.3% *)
P2T	1 + 3c or 1(fast) + 5	29.2%
FAB2	1 + 3a or 1 + 3b	35.8%
F21	1(fast) + 3a + 2 or 1(fast) + 3b + 4	49.0% *)

*) $E_2 < E_1$

The conventional, but tedious way to obtain a model which optimally fits the curves of such a series is to combine all steps being supposed to be active, and then to adapt the activation data and heats of all steps, comparing the experimental plots with computer–generated plots and data from the literature if ever possible. In this way, we have optimized the model POZ7C, considering all seven steps of the scheme (data in Table 2).

Table 2
Activation and signal data [Units: kcal/mole and min] for model POZ7C

step	Ei	lg Ai	−ΔHi
1	9.0	10.7	35
2	16.0	12.9	80
3a	7.0	6.7	45
3b	10.5	10.4	55
3c	11.8	11.5	32
4	21.0	15.6	55
5	12.0	10.6	100

Now it has been shown that, starting with the reference reaction 1 (n = 1) and introducing a further step, respectively, in order to extend the model (occasionally during intermediate neglection of certain steps already considered), the theoretical curve set of the new model (n = n+1 steps) yields an MCC (including all peaks; Table 3) where the introduction of the new step is indicated by the strongest increase of the probability of exactly the corresponding TRM in the distribution list.

Table 3
Mechanistic concentration codes for experiments and three models

	Experiments	Model POZ6
steps	unknown	6
evid.factor	100.0%	62.1%
(S)	−) L) L + + − − 1)	M M) M) + + − 2H −
(M20)	2) −) H) − H) H) − −)	− − − − H H) H +)
(M2)	1L − − − + +) M 1	L + − − 2) + M) 1)

	Model POZ7C	Model POZ7D
steps	7	7
evid.factor	55.2%	58.2%
(S)	L1 +) M − +) − − M2	1 + M + − + H 2
(M20)	1 + − − H) H H H)	H +) − H H H H) H
(M2)	1 + L) − H) H) 2H 2M	1) + L − + + M 2)

Simple rating for comparing elements

Identical elements	1.00
'C' vs. 'C', orders different	0.75
'C' vs. '+)' or '−)' (undistinct)	0.50
Different elements	0.00

Expression for the evidence factor of two strings compared

$$P = \frac{B}{n_{ex} + n_{th} + u}$$

$3 \cdot B$ = rating sum over all elements
n_{ex} = number of intervals exp. string
n_{th} = same, theoretical string
u = number of overlapping intervals

P is maximized by an adequate shift

Progressing from model POZ6 (n = 6; reactions 1,2,3a–c,4; MCC in Table 3) to model POZ7C (n = 7; all reactions; MCC in Table 3), the *new step* 5 in Table 1 is indicated by the fact that the *evidence value* of the corresponding model P2T shows the *strongest increase* of all submodels, namely from 29.2 to 47.2%. On the other hand, the highest absolute percentage, hold by the consecutive model F21* (49.0%), has been lost in favour of model PBA1 (47.9%), a bimolecular reaction competing with an unimolecular one where the reference component is PO_3 (not varied!). If the new distribution is compared with that of an experimental series, involving 6 experiments using $[PO_3]_0$ = 0.05; $[P]_0$ = 0.00, 0.026, 0.05, 0.08, 0.20 and 0.80 M (cf. MCC's in Table 3), the top position of a consecutive type (FAB2 in this case) has been re–gained (47.8%), which also shows the strongest increase of 12%. The other most striking changes in order to reach the experimental evidence values are FAB2* +3, P21* −11, P2T −9, PBA1 −9%. So far, there is no distinct improvement of model POZ6 by the use of model POZ7C. However, these changes imply, in accordance with the scheme and the correlations in Table 2, that a better fit of the data of the (still unfinished) model POZ7C requires the following operations:

1. Decrease of the activation energy of the first reaction (E1) compared with E3c since FAB2 (i.e.E1<E3c), not FAB2* is predominating in the experimental MCC
2. Acceleration of reaction 1 vs. reaction 5
3. Acceleration of reaction 3b, and/or damping of reaction 3c
4. The influences of reactions 3c and/or 5 have to be reduced vs. the concurrent reaction 1.

These proposals of the computer were met by changing for model POZ7C the following activation data, using the same skeleton (Fi = log A of reaction i): E1 −1, F1 +0.1, F3b +0.2, F4 −0.4, F5 −0.2, F3c −0.2, E3b −0.5. Then, the changes of E1 and E3b were recalibrated by further changes of the A–factors of these steps in order to hold the rate coefficients in the centre of the passed temperature interval constant, leading to F1 +1.02, F3b −0.11.

Using now the improved model (= POZ7D) for the simulation of a new theoretical set of curves and determining the new MCC (Table 3), the newly needed changes for the contributions of the TRM's are FAB2 −1, FAB2* 0, P21* +4, P2T 0, PBA1 −1%, which is remarkably less than before: The sum of the five absolute deviations is reduced from 44 (or 40 for POZ6) to 6%. However, the net probability for an agreement of this MCC with the experimental one is only 58.2% (cf. Table 3); the improvement was only concentrated to the five mentioned submodels. Further, one

reason for the bad adaption must be that the scavenging reaction is not exactly observable at linear temperature increase if the initial temperature is below 151 K. Altogether, the model search and the fitting process have as yet not been finished, and POZ6 (62.1%) seems to be even better than POZ7D (58.2%); so one should prefer to modify the skeleton, instead of continuing fitting the activation data at the present stage.

Such comparisons open a fundamental way to examining the quality of an assumed model. However, the evidence values are dependent on the rating matrix and on the procedure needed to consider also empty (i.e. non–correspondent) elements /6/; the latter are missing here, since the four considered MCC's have the same N, namely 8 intervals (elements), which reduces the problems essentially (Table 3).

5. CONCLUSION AND OUTLOOK

The principle of the *self–correcting strategy of reaction analysis* may be summarized as follows: For the study of a system involving unknown kinetics, one starts using the simplest, but mostly wrong assumption of a fitted one–step process. Analyzing the deviations of both the height (or broadness) and the asymmetry of the plot from the primary standardized values leads to a pattern, typical of the pertaining kinetic model. Then, this is compared with patterns of all respective models (here: TRM's) stored in a data bank. The reference character of this strategy makes it sure that unavoidable shortcomings – as e.g. taking a rather arbitrary weight function for the calculation of overall activation data – are partially compensated for; both experimental and theoretical plots are evaluated by the same program.

Application of our pattern–directed program MODEL to numerous reactions in solution has revealed that the best TRM's in general correspond to the steps discussed in the accompanying literature. Further, there are instructive dependencies between the correlation coefficients of the fitted Arrhenius straight line and the corresponding TRM, showing that a good straight line is a very bad criterion for a one–step reaction, compared with the kinetic coordinates /15/. The rather experimental results, mostly obtained using reaction calorimetry and the proven deconvolution formulae for halfwidth and shape index /16/, are now completed by the results of the simulation project, stating that a MCC forms an optimum abstraction vehicle even for any method yielding rate–proportional signals, and may be understood as an alternative of the *reaction matrix*, expanded to practical application. Since the *reaction matrix* represents the shortest mathematical description of any model in homogeneous kinetics /17/, it was proposed by the IUPAC to facilitate the communication between kineticists.

The new tool of pattern recognition (in its widest sense) requires from the chemist to become also familiar with aspects of game theory and decision theory, based on probabilities /18–20/. Problems evolve from the high redundance in the selection of rating functions. In principle, any function monotonously increasing with the utility (outcome, payoff) may be sufficient. Such a function is plausibly offered here by the contribution of coincident experimental and theoretical elements (i.e. equal kinetic behaviour) to all elements of the compared strings. However, the question whether such *evidence values are really absolute probabilities*, also depends on the type of rating matrix and is difficult to answer.

Nevertheless, such *evidence values* as formulated in Table 3 – which should not be confounded with the usually much higher *confidence levels* – are a safe basis for comparing the adequacy of different models. Thus, it was stated that the procedure may be extended to four–reaction models by superposition of two TRM's, one preferred at low, the other at high initial concentration. In the cases of obvious justification of such best models by literature searches, the evidence values of the best single TRM are further increased (up to 90% in favourable cases). There is no reason to assume that the strategy should not be extended to models with more than 4 steps; this is a

question of the development of more effective computers and programs in order to learn ruling the combinatorial variety /5,21/.

An expansion to heterogenkinetic reactions may be discussed if an obligatory general classification concept, involving several (chemical or physical) steps, will have been elaborated; cf. /22,23/. Of course, the abstract kinetic information should be completed by the direct, time–independent view of the participating species and their structures /24/. On the other hand, numerical, especially calorimetric studies have been fundamental in physical chemistry, even for whole chemistry; creating and applying *adequate classification principles* offer the only way to understand something more of the dynamic behaviour of highly complex systems, which agrees with the permanent endeavors of the ICTA to provide for a higher quality of publications dealing with thermoanalytical kinetics /25,26/.

REFERENCES

1 B. Wunderlich and P. Fellner, Thermochim. Acta 110 (1987) 67.
2 J. Fiala, Thermochim. Acta 110 (1987) 11.
3 R. Lutz, in "Informationsentropie oder allmählicher Orientierungsverlust im Datenwald", DECInfo 57 (1989) 15.
4 J. Sestak, J. thermal Anal. 16 (1979) 503.
5 E. Koch, Thermochim. Acta 121 (1987) 253.
6 E. Koch, Thermochim. Acta 152 (1989) 387.
7 E. Koch, Angew. Chem. 85 (1973) 381; Ang. Chem. Int. Ed. Engl. 12 (1973) 381.
8 D. Krug, Thermochim. Acta 10 (1974) 217.
9 M. Balarin, Physics Letters 64A (1978) 435.
10 E. Koch, Theor. Chim. Acta, submitted.
11 A.A. Frost, R.G. Pearson, "Kinetik und Mechanismen chemischer Reaktionen", Verlag Chemie, Weinheim 1964, pp. 81–89.
12 Z.G. Szabo, T. Berces, Z. Phys. Chem. NF57 (1968) 113.
13 E. Koch, "Non–isothermal Reaction Analysis", Academic Press, London 1977, pp. 362–374.
14 R.W. Murray, M.L. Kaplan, J. Am. Chem. Soc. 91 (1965) 5358.
15 E. Koch, J. thermal Anal. 33 (1988) 177.
16 E. Koch, Thermochim. Acta 82 (1984) 293.
17 H. Mauser, "Formale Kinetik", Bertelsmann Univ. Verlag, Düsseldorf, 1974.
18 R. Carnap, W. Stegmüller, "Deduktive Logik und Wahrscheinlichkeit", Springer, Wien, 1959.
19 K.R. Popper, "The Logik of Scientific Discovery", Hutchinson of London, 1974.
20 R.D. Luce, H. Raiffa, "Games and Decisions", Wiley, New York, 1965.
21 E. Koch, Zs. Nagy–Ungvarai, Ber. Bunsenges. Phys. Chem. 91 (1987) 1375.
22 J. Sestak, G. Berggren, Thermochim. Acta 3 (1971) 1.
23 J. Kriz, J. Sestak, Thermochim. Acta 110 (1987) 87.
24 W.F. Hemminger, H.K. Cammenga, "Methoden der Thermischen Analyse", Springer, Berlin, 1989.
25 J.H. Flynn, M. Brown, J. Sestak, Thermochim. Acta 110 (1987) 101.
26 J.H. Flynn, M. Brown, E. Segal, J. Sestak, Thermochim. Acta 148 (1989) 45.