CERAMICS—Silikáty, *35*, s. 1—8 (1991)

Original paper

IMPROVING THE OBJECTIVITY IN THE DETERMINATION OF THE NUMBER OF TERMS IN THE EQUATION OF THE KINETIC MODEL OF A POLYREACTIVE REACTANT

VLADIMÍR KOVÁR[1], IVAN HAVALDA[2]

Faculty of Chemical Technology, Slovak Technical University, Radlinského 9,
812 37 Bratislava

[1]) *Department of the Chemical Technology of Silicates*

[2]) *Department of Chemical Engineering*

Received 7. 5. 1990

The previous paper [2] was concerned with determining the number of terms in the equation modelling a polyreactive reactant on the basis of so-called equalization in the dependence of the minimum sums of deviations squared on the number of terms in the kinetic equation. The present study was aimed at making the method more objective.

The principle is based on transforming the given relationship to one between the estimate of deviations and the number of terms, and on verifying the position of its 'equalization' statistically by means of the F test. It was shown that the kinetic equation with the number of terms established in this way can yield such a dependence of conversion α on time t which reproduces the experimental values within the framework of experimental errors. The equation of the model of a polyreactive reactant has then the form

$$\alpha = \sum_{j=1}^{m} x_j f_j(t),$$

where $x_j \in (0, 1\rangle$ *is the parameter,*
$f_j(t) \in \langle 0, 1\rangle$ *is the growth function,*
m *is the effective number of terms in the equation.*

INTRODUCTION

The course of a reaction of a solid reactant can be successfully described by means of semiempirical relationships based on simplified considerations (models) of the course of the reaction. A special position among these models is occupied by the so-called general (multipurpose) kinetic models with a mathematical background of statistical and probabilistic approach.

The kinetic model of a polyreactive reactant [1], which also belongs to this group of models, is based on the concept of reactivity distribution in the reactant, so that the reactant is regarded as a polyreactive one. On considering such a model on the principle of a reactant comprising several phenomenological proportions with different reactivities (homoreactive fractions), one may express the time dependence of the reactant conversion by the equation

$$\alpha = \sum_{j=1}^{m} x_j f_j(t), \tag{1}$$

where α is the conversion of the reactant as a whole in terms of time t,

x_j is the share of j-th phenomenological homoreactive (PH) fraction in the reactant,

$f_j(t)$ is the concrete growth time function describing the conversion of j-th PH fraction,

m is the number of PH fractions.

The number of terms in the equation and its parameters are determined subsequently from experimental t — α data so as to make the calculated curve describe acceptably the relationship being measured. An acceptable curve is understood to mean such a relationship from which the experimental points do not show deviations exceeding the maximum real measuring error, i.e. $3s$ where s is the standard deviation of the individual measurement (the so-called three-sigma rule commonly used in interval estimates).

In contrast to the requirement of minimum sum of deviations squared which ensures the closest approach of the fit to experimental points, the acceptability criterion is thus oriented toward the reproducibility of the measuring data within the framework of experimental errors. All curves confirming to the acceptability criterion are regarded as being of equal validity in spite of their different sums of deviations squared.

In view of the also known forms of functions $f_j(t)$ ensuring that equation (1) is capable of providing an acceptable description, it can be expected that the given requirement will be met by an equation with various numbers of terms. At the same time, the sum of deviations squared will of course be the smaller the larger the number of terms and thus also the number of parameters of equation (1).

From the standpoint of further utilization of the parameters, it is useful to determine such a number of terms of equation which is necessary for acceptable description of the experimental dependence $\alpha = f(t)$. The minimum required number of terms in equation (1) is called effective.

If the value of s is not known, the effective number of terms is determined from the dependence of the minimum sums of deviations squared (MSDS) on the number of terms m [2]. This relationship shows a monotonously decreasing course for actual values of m, while exhibiting only indistinct changes beyond a certain number of terms ('becoming equalized'). A comparison with the case when the value of s was known showed that equation (1) with this number of terms gave the first acceptable description of the experimental relationship. The number of terms from which the MSDS $= f(m)$ relationship begins to equalize, is then effective and the optimum parameters of equation (1) with this number of terms are accepted as parameters of the model.

The method of determinig the effective number of terms in equation (1) is based on the fact that the ratio of mutually similar numbers is approximately equal to unity. The task of the evaluator is then to compare the neighbouring values of MSDS at a gradually increasing value of m. Effective is that number of terms for which the ratio of two subsequent MSDS values attains, for the first time, a value close to unity.

Although the method is very rapid, it is not always quite unambiguous. It has the disadvantage in that the condition of 'equalization' of the MSDS values can be influenced by a subjective approach of the evaluator.

The present study had the purpose to improve the objectivity in the determination of the effective number of terms in equation (1) for the cases when the standard deviation of the individual measurements is unknown.

SOLUTION AND DISCUSSION

The solution of the problem was sought in the transformation of the MSDS values into mutually comparable statistical quantities. Among these the dispersions σ_m^2 of experimental points around the curves following from equation (1) having the number of terms m are directly associated with the MSDS values. The relationship $\sigma_m^2 = g(m)$ can be expected to be of the same character as the relationship MSDS $= f(m)$, that is that over the range of ,equalizing' exhibited by one of the relationships, the values of the other will likewise show no significant mutual differences. The effective number of terms in equation (1) is then determined by establishing such a number of terms m, beyond which the values of σ_m^2 are already equalized.

Estimates of $\hat{\sigma}_m^2$ of unknown dispersions σ_m^2 are given by the equation [3]

$$\hat{\sigma}_m^2 = \frac{\mathrm{MSDS}_m}{n - p_m}, \tag{2}$$

where n is the number of experimental data pairs t — α,
p_m is the number of parameters in equation (1) with m terms.

As indicated by the given equation, the retelationship $\hat{\sigma}_m^2 = g'(m)$ passes through a minimum in the course of ,equalization' of the numerator when the denominator decreases at the same time. This fact, which is in apparent disagreement with the assumption on the course of the relationship $\sigma_m^2 = g(m)$, is the result of an excessive number of parameters. The lower accuracy of the dispersion estimate is likewise indicative of an unsuitable use of equation (1) with such a number of terms.

Because the insignificant difference in the dispersion estimates is obviously situated in the region of the minimum shown by the relationship $\hat{\sigma}_m^2 = g'(m)$, one can safely consider, with respect to the assumption made above, the value m at the minimum to be either identical with, or very close to, the effective number of terms in equation (1).

The significance of a difference between two dispersions is assessed by means of the F test. The authors have therefore decided to utilize the F test in the determination of the m value, beyond which the relationship MSDS $= f(m)$ or $\sigma_m^2 = g(m)$ becomes equalized.

The F test procedure is based (with respect to the problem in question) on the so-called F distribution of a random variable

$$F_{i,i+1} = \frac{\hat{\sigma}_i^2}{\hat{\sigma}_{i+1}^2} \cdot \frac{\sigma_{i+1}^2}{\sigma_i^2}, \tag{3}$$

where $\hat{\sigma}_i^2$ and $\hat{\sigma}_{i+1}^2$ are estimates of dispersion σ_i^2 and σ_{i+1}^2 respectively.

The distribution of a random variable $F_{i,i+1}$ is described by a function with parameters ν_i and ν_{i+1} — degrees of freedom (see e.g. [3]); the function determines the probability P at which the ratio $F_{i,i+1}$ can acquire a value smaller than the given or chosen value of $F_P(\nu_i, \nu_{i+1})$.

The mutual relations between parameter ν_i, the number of experimental points and the number p_i of parameters in equation (1) having i terms are given by the equation [3]

$$\nu_i = n - p_i. \tag{4}$$

If it holds for the equalization region of the relationship MSDS $= f(m)$ that the dispersions $\sigma_m^2/\sigma_{m+1}^2$ will show insignificant differences (i.e. $\sigma_i^2 \doteq \sigma_{i+1}^2$), then also the random variable

$$F_{i,\,i+1} = \frac{\hat{\sigma}_i^2}{\hat{\sigma}_{i+1}^2} \qquad (5)$$

will exhibit approximate F distribution with parameters ν_i and ν_{i+1}.

In that case, using the F distribution function like in the F test, it is possible to determine some reasonable interval $\langle 0,\ F_P(\nu_i,\ \nu_{i+1})\rangle$ for the $\hat{\sigma}_i^2/\hat{\sigma}_{i+1}^2$ ratio, in which the latter will fit with a very high probability P (if $\hat{\sigma}_i^2 < \hat{\sigma}_{i+1}^2$, it is obviously the interval $\langle 0,\ F_P(\nu_{i+1},\ \nu_i)\rangle$). In technical practice, the top limit of this interval, also called the critical value, is chosen usually for a probability of 99% or 95%.

If the value of the ratio $\hat{\sigma}_i^2/\hat{\sigma}_{i+1}^2$ does not fit into such an interval, that is when

$$F_{i,\,i+1} = \frac{\hat{\sigma}_i^2}{\hat{\sigma}_{i+1}^2} > F_P(\nu_i,\ \nu_{i+1}) \qquad (6)$$

holds with a risk smaller than 1 — P, the ratio $\hat{\sigma}_i^2/\hat{\sigma}_{i+1}^2$ can be regarded as not exhibiting the F distribution and the values of dispersions σ_i^2 and σ_{i+1}^2 can be taken to show a significant difference.

The conclusions derived were verified on a 24-point model set of t-α data from a previous study [2] where the effective number of terms in equation (1) was found to be equal to three and the standard deviations of the individual measurements were known. The dependence of MSDS on the number of terms m and the number of parameters p_m was as follows [2] (also refer to Fig. 1):

m	1	2	3	4
p_m	1	3	5	7
MSDS $\times 10^3$	213.677	8.807	1.253	0.526

The ratios of two subsequent MSDS, which should approach unity, were therefore 24.3, 7.0 and 2.4. It is thus obvious that the determination of an equal effective number of terms ($m = 3$) on the basis of these values only, was not unambiguous.

The values of $\hat{\sigma}_m^2$ supplemented by calculation for equation (1) having the number of terms m are as follows (also refer to Fig. 1):

m	1	2	3	4
$\hat{\sigma}_m^2 \times 10^4$	92.903	4.194	0.660	0.309

and their mutual ratios ($F_{i,i+1}$) are as follows:

$$F_{1,2} = 22.1518 \quad F_{2,3} = 6.3575 \quad F_{3,4} = 2.1329$$

On selecting 0.99 and 0.95 as the respective values of P, the critical values obtained are [4]

$$F_{0.99}(23,21) = 2.8183 \quad F_{0.99}(21,19) = 2.9810 \quad F_{0.99}(19,17) = 3.1857$$
$$F_{0.95}(23,21) = 2.0633 \quad F_{0.95}(21,19) = 2.1438 \quad F_{0.95}(19,17) = 2.2429$$

The results given above indicate that the dispersions σ_1^2, σ_2^2 and σ_3^2 show significant mutual differences, while the conclusion cannot be accepted for the pair σ_3^2 and σ_4^2 with an error risk smaller than be 1%. In this way, the ,equalizing‘ of the MSDS$_m$ and σ_m^2 values beyond $m = 3$ can thus be regarded as being objectively proved.

To elucidate the mutual relations between the effective number of terms in equation (1) and the value of m for which the relationship $\hat{\sigma}_m^2 = g'(m)$ acquires a minimum, the results of the previous study [2] were supplemented by a calculation using five terms in equation (1). For function $f_j(r)$ in the standard form

$$f_j(t) = 1 - \exp(-r_j t), \tag{7}$$

where r_j is the parameter expressing the reactivity of j-th homoreactive fraction, the following results were obtained:

Optimum parameters

x_j	0.046	0.114	0.197	0.290	0.354
r_j/min^{-1}	5.95	0.34	0.46×10^{-1}	1.03×10^{-2}	0.49×10^{-2}

the minimum sum of deviations squared, MSDS$_5$ was 0.499×10^{-3}, the dispersion estimate $\hat{\sigma}_5^2 = 0.333 \times 10^{-4}$.

Compared to the value of MSDS$_4$ ($= 0.526 \times 10^{-3}$), that of MSDS$_5$ remained virtually unchanged (MSDS$_4$: MSDS$_5$ = 1.05) and one can safely expect the values of MSDS$_m$ to change even less further on. At the same value of MSDS$_m$ and a decreasing value of the denominator in equation (2) resulting from the increase in the number of parameters p_m, the estimate $\hat{\sigma}_m^2$ of the dispersion will therefore keep rising. The value of the dispersion estimate at the number of terms $m = 4$ (close to the effective number) is thus the minimum one.

On considering the relationship $\sigma_m^2 = g(m)$ 'equalized' from $m = 3$ upwards, the values of the pair σ_4^2 and σ_5^2 must also show an insignificant difference. For $P = 0.99$ and $P = 0.95$ it holds that

$$F_{5,4} = 1.0753$$
$$F_{0.95}(15,17) = 2.3077 \quad F_{0.99}(15,17) = 3.3117$$

from which it follows that the dispersions σ_4^2 and σ_5^2 can actually be regarded as being approximately identical.

In addition to this, not only the neighbouring but also any two arbitrary pairs of dispersion should not exhibit any significant differences in the ,equalized‘ portion of relationship $\sigma_m^2 = g(m)$. The values available allow this assumption to be verified on the pair of σ_3^2 and σ_5^2 for which it holds that

$$F_{3,5} = 1.9835$$

$$F_{0.95}(19,15) = 2.3398 \quad F_{0.99}(19,15) = 3.3961$$

The values of σ_3^2 and σ_5^2 are thus also approximately identical.

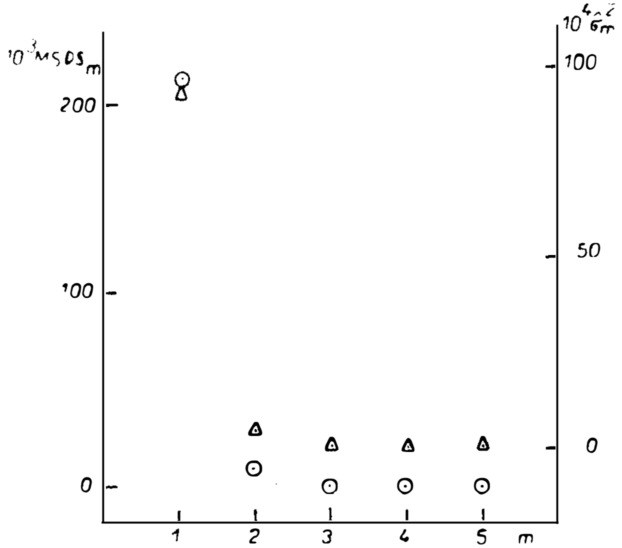

Fig. 1. Minimum sums of deviations squared MSDS$_m$ and estimates $\hat{\sigma}_m^2$ of deviations of experimental points vs.the number m of terms in equation (1).
○ — MSDS$_m$, △ — $\hat{\sigma}_m^2$.

CONCLUSION

The objectivized method for determinig the effective number of terms in equation (1) can be summarized as follows:

At a gradually increasing number m of terms in equation (1), a suitable optimizing procedure is used to determine the parameters of the equation. For each value of m, the MSDS is calculated and using equation (2), the dispersion estimate $\hat{\sigma}_m^2$ for the experimental points around the curve given by equation (1) with m terms is established. The ratio of the last subsequent values of dispersion estimates is then compared with the tabulated critical value $F_P(\nu_i, \nu_{i+1})$, where P is a value very close to unity and parameters ν_i and ν_{i+1} are calculated by means of equation (4).

If the condition (6) is met, the relationships MSDS$_m = f(m)$ and $\sigma_m^2 = g(m)$ are regarded as being still unequalized. The calculation then proceeds to the next values of m until the conditon (6) is not met for the first time.

In the case of opposite inequality (6), that is when there are not sufficient reasons for rejecting the hypothesis on the insignificant difference between dispersions σ_i^2 and σ_{i+1}^2, the relationships MSDS $= f(m)$ and/or $\sigma_m^2 = g(m)$ are regarded as being ‚equalized', and the number of terms i, from which this equalization takes place, as being effective.

References

[1] Jesenák V.: *Proc. VIth National Conference on Thermal Analysis 'Termanal' 73*, ES SVŠT Bratislava 1973, p.37–47.
[2] Kovár V., Bachan J.: Ceramics—Silikáty 34, 289 (1990).
[3] Anděl J.: *Mathematical Statistics* (in Czech), SNTL Prague 1978.
[4] Likeš J., Laga J.: *Elementary Statistical Tables* (in Czech), SNTL, Prague 1978.

OBJEKTIVIZÁCIA URČENIA EFEKTÍVNEHO POČTU ČLENOV ROVNICE KINETICKÉHO MODELU POLYREAKTÍVNEHO REAKTANTA

Vladimír Kovár, Ivan Havalda

Chemickotechnologická fakulta SVŠT, Radlinského 9, 812 37 Bratislava

Na popis závislosti konverzie α od času t pri odreagovávaní tuhého reaktanta sa úspešne používa aj kinetický model polyreaktívneho reaktanta, podľa ktorého

$$\alpha = \sum_{j=1}^{m} x_j f_j(t), \tag{1}$$

kde $x_j \in (0, 1\rangle$ je parameter,
$f_j(t) \in \langle 0, 1\rangle$ — konkrétna funkcia majúca vzťah k reakcii,
m — neurčený počet členov rovnice,

Počet členov rovnice a jej parametre sa určujú dodatočne z experimentálnych údajov. Rovnica reprodukuje experiment tým lepšie, čím má viac členov. Počet členov, od ktorého sa reprodukovateľnosť už významne nemení sa označuje za efektívny. Tento počet sa určuje zo závislosti rozptylov σ_m^2 od počtu členov m, pričom odhad $\hat{\sigma}_m^2$ rozptylu je daný vzťahom

$$\hat{\sigma}_m^2 = \frac{\mathrm{MSDS}_m}{n - p_m}, \tag{2}$$

kde n je počet experimentálnych dvojíc údajov t — α,
p_m — počet parametrov rovnice (1) s počtom členov m.

Metodika určenia efektívneho počtu členov rovnice (1) spočíva v tom, že pri postupne sa zväčšujúcom počte členov sa F-testom testuje významnosť odlišnosti dvoch po sebe nasledujúcich rozptylov. V prípade, že sa rozptyl významne nelíši od predchádzajúceho rozptylu, považuje sa predchádzajúci počet členov rovnice za efektívny.

Obr. 1. Závislosť minimálnych súčtov štvorcov odchýlok MSDS_m, *resp. odhadov* $\hat{\sigma}_m^2$ *rozptylov experimentálnych bodov od počtu m členov rovnice (1)* ○ — MSDS_m, △ — $\hat{\sigma}_m^2$.

V. Kovár, I. Havalda:

ОБЪЕКТИВИЗАЦИЯ УСТАНОВЛЕНИЯ ЭФФЕКТИВНОГО КОЛИЧЕСТВА ЧЛЕНОВ УРАВНЕНИЯ КИНЕТИЧЕСКОЙ МОДЕЛИ ПОЛИРЕАКТИВНОГО РЕАКТАНТА

Владимир Ковар, Иван Гавалда

Химико-технологический факультет Словацкого политехнического института,
Радлинского 9, 812 37 Братислава

Для описания зависимости конверсии α от времени t при прореагировании твердого реактанта с успехом используется даже кинетическая модель полиреактивного реактанта, согласно которой

$$\alpha = \sum_{j=1}^{m} x_j f_j(t), \qquad (1)$$

где $x_j \in (0,1\rangle$ — параметр, $f_j(t) \in \langle 0,1\rangle$ — конкретная функция, относящаяся к реакции, m — неустановленное количество членов уравнения.

Количество членов уравнения и его параметры устанавливаются добавочно, на основании экспериментальных данных. Уравнение воспроизводит эксперимент тем лучше, чем больше имеет членов. Количество членов, с которого воспроизводимость значительно не изменяется, называется эффективным. Данное количество определяется на основании зависимости рассеяний, σ_m^2 от количества членов m, причем оценка $\hat{\sigma}_m^2$ рассеяния дана отношением

$$\hat{\sigma}_m^2 = \frac{\mathrm{MSDS}_m}{n - p_m}, \qquad (2)$$

где n — количество экспериментальных пар данных t — α, p_m — количество параметров уравнения (1) с количеством членов m.

Методика установления эффективного количества членов уравнения (1) заключается в том, что при постепенно увеличивающемся количестве членов с помощью F-теста устанавливается величина значения двух последующих друг за другом рассеяний. В случае, что рассеяние значительно не отличается от предыдущего рассеяния, считается предшествующее количество членов уравнения эффективным.

Рис. 1. Зависимость минимальных сумм квадратиков отклонений MSDS_m, *или оценок* $\hat{\sigma}_m^2$ *рассеяний экспериментальных точек от количества* m *членов уравнения (1);* ○ — MSDS_m, △ — $\hat{\sigma}_m^2$.