# MODELLING OF CATALYTIC ACTIVITY OF TRANSITION METALS IN HYDROGENOLYSIS OF ETHANE BY THE PATTERN RECOGNITION APPROACH

Oldřich Štrouf[a], Jiří Fusek[a] and Karel Kuchynka[b]

[a] *Institute of Inorganic Chemistry, Czechoslovak Academy of Sciences, 250 68 Řež and*
[b] *J. Heyrovský Institute of Physical Chemistry and Electrochemistry, Czechoslovak Academy of Sciences, 121 38 Prague*

Received January 21st, 1980

For modelling the catalytic activity in hydrogenolysis of ethane, thirty transition metals were characterized by a set of fifteen fundamental physical variables for each of them. The dimensionality reduction of this physical model up to the dimensionality seven was carried out by the determination of seven linearly almost independent variables. This set was used for the definition of the class of catalytically active metals on the basis of their similarity using an asymmetric (one-class) pattern recognition approach. The values of similarity measure were approximated by the polynomial function of the second degree which is accounted for as a mathematical representation of the structure of the active class.

The pattern recognition approach was developed for classification of the objects of a system by an analysis of their multivariate data in the term of similarity[1]. Different measures were formulated for a quantitative evaluation of the similarity, on the basis of them various models can be constructed. Some of these models work with *a priori* knowledge about the structure of classes, whereas the remaining ones make no presumption about the structure.

The latter approach has been used also in our pattern recognition method[2] where the measure of similarity, $S_q$, is the Euclidean distance in the measurement space which is transformed and normalized for each class of the system separately and independently. This disjoint method has been recently applied for the search for regularities in chemical structures of complex hydrides of the 3rd main group elements* determining the stability of these compounds[2]. The efficiacy of the method has been tested by the direct comparison of the results with those obtained previously by highly sophisticated SIMCA method[3].

The method is applied here for an asymmetric[4,5] pattern recognition analysis of catalytic activity of some transition metals in the hydrogenolysis of ethane:

* The denomination IIIB was used in refs[2,3]. More frequently used denomination IIIA was accepted in ref.[5] and it will be used in all forthcoming communications.

$$C_2H_6 + H_2 \xrightarrow{\text{metal}} 2\,CH_4 \qquad (A)$$

The metals are characterized by fifteen fundamental physical parameters available for all 30 metals included into the study. These data are significantly more exact in comparison with those of complex hydrides[6,7] which contain also some derived data and structural descriptors. Moreover, for eleven metals the levels of activity are given quantitatively in terms of the rate constants[8], whereas the stability of the complex hydrides can be expressed only qualitatively. Therefore, it is possible to define the class(es) more exactly in the former case than in the latter one.

On the basis of the above data, the values of the measure of similarity $S_q$ (here $S_a$) for the metals were estimated by the mentioned method[2] and then approximated by a polynomial function, the coefficients of which were determined by the simplex procedure given in the preceding communication[9]. Two main goals of such a mathematical approximation are evident: the practical goal is an easy calculation of interpolated values of the measure of similarity for the metals with unknown activity and the fundamental one is a mathematical representation of class structure important for the generalization of system behaviour.

## EXPERIMENTAL

*Data.* Values of 15 variables available for the physico-chemical model under study were taken from three sources — those of the ionization potentials from ref.[10], the Debye's temperatures from ref.[11] and all other from the Sargent–Welch table[12]. The value of heat of vaporization of Nb (*c.* $3{\cdot}98 \,.\, 10^{-2}$ J kg$^{-1}$) was taken from ref.[13] and of Debye's temperature of Sc (360 K) from ref.[14]. For pattern recognition purposes the data were autoscaled. The catalytic activity of eleven metals is given by Sinfelt[8].

*Computation.* Pattern recognition computation of similarity was carried out by the Gier-Algol-III program at the Computer Centre, Institute of Nuclear Research, Řež. The simplex optimization of polynomial approximation of the measure of similarity was performed on a Hewlett–Packard 9825 type computer.

## RESULTS AND DISCUSSION

### *Dimensionality Reduction of the Model*

The reduction of the initial dimensionality $D = 15$ was achieved by the method described elsewhere[15]. The loss of information in the relative eigenvalue scale is given in Table I and Fig. 1. Variables whose deletion causes only a loss lower than 2·2% $\lambda$ were discarded (the variables $x_i$ with $i = 8-15$ in Table I). The set of remaining seven linearly almost independent variables $x_i$ $(i = 1-7)$ seems to be sufficient for characterization of the active class of the modelled system as is supported by the changes of the measure of similarity $S_a$ for the individual metals during the stepwise dimensionality reduction process. The change of the active class volume

as defined by the maximum value of $S_a$ ($S_a^{max}$ in Table I), as well as by the change of the number of metals with unknown activity classified as active ones (Table I)

TABLE I

Dimensionality Reduction of the Model of Transition Metals Catalytic Activity in Hydrogenolysis of Ethane

| $D^a$ | Variable $x_i$ | % $\lambda^b$ | $S_a^{max}$ | $N^c$ |
|---|---|---|---|---|
| 1 | molar heat capacity | — | — | — |
| 2 | covalent radius | 43·8 | 5·29 | 18 |
| 3 | electric conductivity | 24·6 | 3·54 | 16 |
| 4 | heat of fusion | 14·8 | 2·71 | 10 |
| 5 | 1st ionization potential | 11·2 | 2·18 | 5 |
| 6 | specific heat | 4·6 | 1·83 | 4 |
| 7 | 3rd ionization potential | 2·2 | 1·57 | 3 |
| 8 | Debye's temperature | 1·5 | 1·37 | 3 |
| 9 | electronegativity | 0·78 | 1·37 | 2 |
| 10 | 2nd ionization potential | 0·45 | 1·22 | 1 |
| 11 | atomic volume | 0·30 | 1·22 | 1 |
| 12 | heat of vaporization | 0·15 | 1·22 | 1 |
| 13 | molar electric conductivity | 0·10 | 1·22 | 1 |
| 14 | melting point | 0·07 | 1·10 | 1 |
| 15 | atomic radius | 0·05 | 1·10 | 1 |

[a] Dimensionality of the model (see ref.[15]); [b] relative loss of information (see ref.[15]); [c] number of the metals of unknown catalytic activity[8] classified at given dimensionality $D$ as active metals.

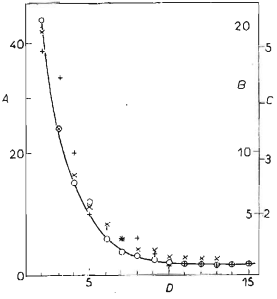


FIG. 1

Dimensionality Reduction in Modelling of Catalytic Activity of Transition Metals in Hydrogenolysis of Ethane

$A$ ○ is a relative loss of information % $\lambda$ (Table I). $B$ + is the number of metals with unknown catalytic activity classified at the given dimensionality $D$ as active metal (Table I). $C$ × is a volume of the active class defined by the values of $S_a^{max}$ at the given dimensionality $D$ (Table I).

fit the curve of information loss very closely (Fig. 1). The latter change limits to a practically complete "expansion" of the active class over the total set of the studied 30 metals. The activity level calculation described in the forthcoming paper[16] confirms also the suggested definition of the class of catalytically active metals.

*Classification of Metals*

For the classification of the metals as catalytically active in the hydrogenolysis of ethane a recently described[2] measure of similarity was used. Eleven transition metals known to be active[8] were treated as a training set for the active class definition. No information about the activity is available in ref.[8] for the remaining 19 metals studied here. This is the reason for using "asymmetric"[4,5], one-class pattern recognition approach. The $S_a^{max}$ found in the training was accounted for as a boundary of the active class volume. In the studied case three metals, *viz.*, W, Mo and Mn with unknown activity possess $S_a$ lower than $S_a^{max}$ and, therefore, they were included additionally into the active class (Table II). Thus the minimum set of fourteen active

TABLE II

Values of Measure of Similarity ($S_a$) and Their Approximations ($S'_a$) for Transition Metals in Seven-dimensional Model of Catalytic Activity in Hydrogenolysis of Ethane

| Group B | Metal | $S_a$ | $S'_a$ | Group B | Metal | $S_a$ | $S'_a$ |
|---|---|---|---|---|---|---|---|
| I | Cu[b] | 1·42 | 1·42 | | Ta | 22·2 | — |
| | Ag | 5·48 | — | VI | Cr | 4·59 | — |
| | Au | 3·72 | — | | Mo[a] | 0·72 | 0·93 |
| II | Zn | 6·38 | — | | W[a] | 0·79 | 0·62 |
| | Cd | 12·3 | — | VII | Mn[a] | 1·52 | 1·46 |
| | Hg | 10·6 | — | | Re[b] | 0·74 | 0·90 |
| III | Sc | 10·5 | — | VIII | Fe[b] | 1·26 | 1·18 |
| | Y | 35·2 | — | | Ru[b] | 0·98 | 0·73 |
| | La | 76·0 | — | | Os[b] | 0·47 | 0·60 |
| | Ce | 51·6 | — | | Co[b] | 0·70 | 0·89 |
| IV | Ti | 5·32 | — | | Rh[b] | 0·48 | 0·42 |
| | Zr | 12·3 | — | | Ir[b] | 1·25 | 1·18 |
| | Hf | 37·2 | — | | Ni[b] | 1·07 | 1·09 |
| V | V | 3·57 | — | | Pd[b] | 1·57[c] | 1·52 |
| | Nb | 2·67 | — | | Pt[b] | 1·25 | 1·29 |

[a] Metals with $S_a < S_a^{max}$ the activities of them are not given in ref.[8]; [b] metals with measured activity[8] used here for the definition of active class by training; [c] $S_a^{max}$.

metals was finally formed. This set was used in the polynomial approximation of the measure of similarity $S'_a$.

The above mentioned relative changes of the similarity measure $S_a$ caused by the stepwise deletion of variables during the dimensionality reduction exhibit an interesting behaviour; the measure $S_a$ for most metals changed remarkably at the $D = 8 \rightarrow 7$ dimensionality reduction step for which the information loss is higher than 2·2% $\lambda$ (Table I, Fig. 1) or even at lower dimensionalities. Only for Rh and Nb such a significant change was found at the step $D = 10 \rightarrow 9$ and, thus, the values of $S_a$ at $D = 7$ for these metals are rather dubious. Nevertheless, the correctness of the above dimensionality reduction from $D = 15$ up to $D = 7$ seems to be sufficiently supported by the character of the $S_a$ changes in all remaining 28 cases.

*Mathematical Approximation of Similarity Measure*

First, the linear polynom (*I*) $a_i \neq 0$, $b_i = 0$ was tested for an approximation of the similarity measure $S'_a$ of the active metals class:

$$S'_a = \sum_{i=1}^{7} a_i x_i + \sum_{i=1}^{7} b_i x_i^2 + c\,, \qquad (I)$$

where $0 \leqq S'_a \leqq 1{\cdot}57$. The set of coefficients $a_i$ for seven linearly almost independent variables $x_i$, $i = 1-7$ (Table I) were adjusted by the simplex procedure described in the preceding paper[9]. In the simplex optimization the responses $r = \sum_{k=1}^{14}(S_a - S'_a)^2$ were minimized for fourteen active metals[8] whose measures of similar-

TABLE III
Coefficients of the Equation (*I*) ($N = 7$, $c = -12{\cdot}142$)

| Variable[a] | Coefficient $a_i$ | Coefficient $b_i$ |
|---|---|---|
| 1 | 1·045 | $-1{\cdot}355 \cdot 10^{-1}$ |
| 2 | 3·289 | 3·605 |
| 3 | −6·994 | $1{\cdot}049 \cdot 10^{1}$ |
| 4 | $-1{\cdot}072 \cdot 10^{-1}$ | $3{\cdot}067 \cdot 10^{-2}$ |
| 5 | $-5{\cdot}276 \cdot 10^{-2}$ | $2{\cdot}446 \cdot 10^{-4}$ |
| 6 | $2{\cdot}419 \cdot 10^{1}$ | $5{\cdot}227 \cdot 10^{1}$ |
| 7 | $7{\cdot}803 \cdot 10^{-4}$ | $2{\cdot}923 \cdot 10^{-6}$ |

[a] See Table I.

ity in the above pattern recognition classification were $S_a \leq 1{\cdot}57$ (Table II). Nevertheless, this optimization yields only a poor accordance of the values of $S_a$ and $S'_a$ as well as a rather high value of the best minimum response $r^B$. This indicates that the approximation by the polynomial function of the first degree is not sufficient for the representation of the active class structure.

Therefore, the polynom of the second degree (Eq. (*1*), $a_i$, $b_i \neq 0$) was tested, the coefficients $a_i$ and $b_i$ (Table III) being optimized by our simplex procedure[9], too. The values of $S'_a$ are given in Table II. The final best response $r^B$ was 0·237.

*Structure of Classes*

As shown above, the structure of the active class can be mathematically represented by the polynom of the second degree which gives an acceptable agreement of the $S_a$ and $S'_a$ values (Table II). Nevertheless an extrapolation beyond the boundary value of the active metals class ($S_a = 1{\cdot}57$) is not valid as is indicated by insufficient accordance between calculated and approximated values of the measure of similarity. This shows that the structure of "inactive" class (metals with $S_a < 1{\cdot}57$) is different of that of the active class. This fact is supported additionally by a relatively higher best response $r^B$ resulting after an exhaustive simplex optimization made over the original set of thirty nonclassified metals.

## CONCLUSIONS

The similarities of the studied transition metals with respect to their catalytic activity in the hydrogenolysis of ethane were estimated by the one-class asymmetric pattern recognition method on the basis of their seven linearly independent variables of basic physical importance. The values of the measure of similarity were approximated by a polynom of the second degree with the coefficients optimized by the simplex procedure, the approximation being invalid for the "inactive" class. Such couple of steps of the chemical systems modelling – the pattern recognition classification by means of a disjoint method and the *a posteriori* class structure estimation – can be highly suitable especially for the modelling of structurally inhomogeneous systems as it was demonstrated here.

### REFERENCES

1. Wold S., Sjöström M. in the book: *Chemometrics. Theory and Application* (B. R. Kowalski, Ed.), p. 243. ACS Symposium Ser. 52. American Chemical Society, Washington 1977.
2. Fusek J., Štrouf O.: This Journal *44*, 1362 (1979).
3. Štrouf O., Wold S.: Acta Chem. Scand. *A31*, 391 (1977).
4. Albano C., Dunn W., III, Edlund U., Johansson E., Nordén B., Sjöström M., Wold S.: Anal. Chim. Acta Comput. Optim. *103*, 429 (1978).
5. Wold S., Štrouf O.: Acta Chem. Scand. *A33*, 463 (1979).

6. Štrouf O., Wold S.: *Data Base I. Stability of Complex Hydrides (MODEL-I)*. Umeå University, Umeå 1977.
7. Wold S., Štrouf O.: *Data Base II. Stability of Complex Hydrides*. Umeå University, Umeå 1978.
8. Sinfelt J. H.: Catal. Rev. *3*, 175 (1969).
9. Štrouf O., Fusek J.: This Journal, in press.
10. Moore C. E.: *Atomic Energy Levels*, Vol. III, p. 34. Nat. Bureau of Standards, Washington 1958.
11. *Fiziko-khimicheskiie Svoistva Elementov* (G. V. Samsonov, Ed.), p. 242. Naukova Dumka, Kiev 1965.
12. *Periodic Table of the Elements*. Sargent-Welch 1968.
13. Brown D. in the book: *Comprehensive Inorganic Chemistry* (J. C. Bailar, jr, M. J. Emeléus, R. Nyholm, A. F. Trotman-Dickenson, Eds), Vol. III, p. 559. Pergamon Press, Oxford 1973.
14. Kittel C.: *Introduction to Solid State Physics*, Vol. IV, p. 219. Wiley, New York 1958.
15. Štrouf O., Fusek J.: This Journal *44*, 1370 (1979).
16. Kuchynka K., Fusek J., Štrouf O.: This Journal, in press.

Translated by the author (O. Š.).