An Improved Procedure for the Description of Substituent Effects on Equilibrium and Rate Constants, based on Factor Analysis of the Whole Body of Appropriate Data

By **Günther H E. Nieuwdorp** * and **C. Leo de Ligny,** * Laboratory for Analytical Chemistry, University of Utrecht, Utrecht, The Netherlands
Hans C. van Houwelingen, Institute for Mathematical Statistics, University of Utrecht, Utrecht, The Netherlands

By comparing the two most widely used descriptions of substituent effects by Hammett type substituent constants, *viz.*, the Taft equation and the Yukawa–Tsuno relation, it is shown that a general descriptive equation should be of the type: $\log (K_X/K_H) = \sum_{i=1}^{n} \rho_i\sigma_i$. In this equation, the left hand side is the substituent effect, ρ is a reaction constant, and σ is a substituent constant. With the aim of finding optimal values of ρ and σ which can be used in this general equation, a large number (570) of $\log (K_X/K_H)$ values for chemical reactions and equilibria was selected from the literature. These data were taken for 76 reactions or equilibria, and 17 substituents. A special statistical procedure, which included factor analysis, was designed and applied to extract optimal ρ and σ values from the data. The results of the procedure show that only three substituent constants, σ_I, σ_R, and σ_E, are required to describe $\log (K_X/K_H)$ values for all reaction types mentioned by Ehrenson, Brownlee, and Taft, *viz.* σ_I, σ_R^0, σ_R(BA), σ_R^+, and σ_R^- reaction types. So, the general regression equation is $\log (K_X/K_H) = \rho_I\sigma_I + \rho_R\sigma_R + \rho_E\sigma_E$. The constants σ_I and σ_R are comparable to Taft's σ_I and σ_R^0. The magnitude of the $\rho_E\sigma_E$ term is indicative of the reaction type: ρ_E is zero for σ_R^0 reactions, negative for *para*-substituted σ_R^- reactions, and positive for σ_R^+ reactions (the substituent effects were defined in such a way that ρ_I and ρ_R are positive for all reactions). Optimal values of σ_I, σ_R, and σ_E are given for the 17 substituents. As an illustration of their applicability, regression analysis of phase equilibrium data on σ_I, σ_R, and σ_E is performed, and the results are compared with those of regression analysis on the Taft σ parameters.

CONSIDER a chemical reaction or equilibrium of the type $XNY + R \rightleftharpoons XNZ$ where X is a variable substituent, N is a common carbon skeleton, Y and Z are fixed reactive centres in the molecules, and R is a common reactant. For the case that N is a benzene ring the effect of X on the reaction rate or equilibrium constant K was originally formulated by Hammett [1] as in equation (1). The subscripts X and H refer to the substituted and unsubstituted compounds, respectively,

$$\log (K_X/K_H) = \rho\sigma \qquad (1)$$

ρ is a reaction dependent parameter, and σ is a substituent dependent parameter. Hammett chose as a reference reaction series ($\rho = 1$) the dissociation of substituted benzoic acids in water at 25 °C. Not only the nature of the substituent determines its σ value but also its position in the benzene ring relative to the reactive centre. Therefore each substituent has a σ_m (*meta*) and a σ_p (*para*) value, mainly owing to the different resonance interaction possibilities in these positions. In the case of *ortho*-substituents steric and sometimes hydrogen bonding effects add to the 'electronic' (*i.e.* inductive and resonance) interactions of the substituents. No general description of data for *ortho*-substituted systems could be developed. It was soon recognized that the Hammett relation could describe substituent effects in only a relatively small area of types of reactions or equilibria. To widen the scope of applications the relation was extended in various ways during the last two decades. Two of the most frequently used extensions will be shortly described here: The dual parameter equation of Taft [2,3] and the Yukawa–Tsuno [4] relation.†,‡

The Taft equation is of the form (2). In this equation the subscripts I and R refer to the inductive effect and the resonance effect of a substituent, respectively. For each substituent only one σ_I value (derived from aliphatic reaction series) is given. The σ_R value depends on the type of reaction under investigation.

$$\log (K_X/K_H) = \rho_I\sigma_I + \rho_R\sigma_R \qquad (2)$$

In a review by Ehrenson *et al.*[3] four types of reactions, and thus four sets of σ_R values, were proposed: σ_R^- type reactions: direct (through, exalted) resonance is possible between an electron-withdrawing substituent and an electron-donating reaction centre [the classical example is the dissociation equilibrium of the *p*-nitroanilinium ion ($X = NO_2$, $Y = NH_2$, $R = H^+$, $Z = NH_3^+$; direct resonance is possible between NO_2 and NH_2)]; σ_R^+ type reactions: direct resonance is possible between an electron-donating substituent and an electron-withdrawing reaction centre (the reference reaction is the solvolysis of *para*-substituted t-cumyl chlorides); σ_R^0 type reactions: direct resonance is considered to be impossible (the dissociation of substituted phenylacetic acids belongs to this class); σ_R(BA) type reactions: reactions of substituted benzoic acids or reactions that are considered to be similar to these. Ehrenson *et al.*[3] give a fifth set of resonance constants, *i.e.* σ_R^-(P) values, for phenolic type reactions, but they do not recommend these for general use.

$$\log (K_X/K_H) = \rho\,(\sigma_{m,p}^0 + r^{\pm}\Delta\sigma_R^{\pm}) \qquad (3)$$

The Yukawa–Tsuno relation is of the form (3). The index $\pm$ refers to σ_R^+ or σ_R^- type reactions as defined

---

† For details and theoretical aspects of the derivation of these relations see refs. 3 and 5.

‡ Other extensions as the Swain–Lupton equation [6a] and Dewar's FMMF treatment [6b] will not be discussed here because these approaches cannot describe the effect of direct resonance between a substituent and reactive centre.

above. For electron-withdrawing substituents $\Delta\sigma_R^+$ values are zero and $\Delta\sigma_R^-$ values are positive. For electron-donating substituents $\Delta\sigma_R^+$ values are negative and $\Delta\sigma_R^-$ values are zero. As the authors assume σ_m^0 and σ_p^0 to be linear combinations of σ_I and σ_R^0, the Yukawa–Tsuno relation can be expressed as a four parameter extension of the Taft equation, holding for the five reaction types mentioned above [equation (4)].*

$$\log (K_X/K_H) = \rho_I\sigma_I + \rho_R\sigma_R^0 + r^+\Delta\sigma_R^+ + r^-\Delta\sigma_R^- \quad (4)$$

The fundamental restrictions inherent in the use of the Taft equation or the Yukawa–Tsuno relation can best be compared on the basis of equation (4).

(A) In the Yukawa–Tsuno approach the parameters ρ_I and ρ_R in equation (4) are subject to the constraints (5)—(7) where c, c', and c'' are constants. According to

$$\rho_{I,m}/\rho_{R,m} = c \quad (5)$$

$$\rho_{I,p}/\rho_{R,p} = c' \quad (6)$$

$$\rho_{I,m}/\rho_{I,p} = c'' \quad (7)$$

Taft, these constraints do not apply.

(B) The use of $\sigma_R^+(\sigma_R^-)$ parameters in the Taft equation implies a constant ratio $r^+/\rho_R(r^-/\rho_R)$ in equation (4) for all $\sigma_R^+(\sigma_R^-)$ type reaction series. In the Yukawa–Tsuno approach this ratio is variable.

(C) For all *meta*-substituted reaction series, Yukawa and Tsuno assume $r^+ = r^- = 0$. This is not stipulated by Taft.

(D) For electron-withdrawing substituents the Yukawa–Tsuno approach prescribes $\Delta\sigma_R^+ = 0$. An identical restriction is inherent in the Taft approach, as can be judged from the equality of the proposed values of σ_R^0 and σ_R^+ for these substituents.[3] For electron-donating substituents Yukawa and Tsuno analogously assume $\Delta\sigma_R^- = 0$. From the proposed σ_R^0 and σ_R^- values for the Taft equation [3] it appears that this restriction is effectively valid only for some of these substituents.

Most of the restrictions (1)—(4) have been subject to doubt and discussion. Even the application of the well accepted σ_R^0 set of parameters has its hazards, as was pointed out by Hoefnagel and Webster.[7] In the following only points (C) and (D), mentioned above, will be discussed as examples of dubious restrictions.

(C). From the classical concept of direct resonance the assumption that *meta*-substituents cannot give direct resonance with a reaction centre seems correct. However, the use of only one type of substituent parameter for all types of reactions, *i.e.* σ_m, as is prescribed for the Yukawa–Tsuno relation, seems not justified in view of some experimental data: a plot of log (K_X/K_H) values for the dissociation of *meta*-substituted anilinium ions [8] (for *para*-substituents the classical σ_R^- type reaction) *versus* corresponding data for the solvolysis of t-cumyl chlorides [9] (the classical σ_R^+ type reaction) shows considerable scatter. Ehrenson *et al.*[3] state that the assumption, that all *meta*-substituted reaction series are σ_R^0 type reactions, is not justified, statistically.

(D). The assumption that for electron-withdrawing substituents $\Delta\sigma_R^+ = 0$ is a result of the concept that direct resonance can operate only in one way (*i.e.* $r > 0$). This means that an eventual decrease of an electron-withdrawing substituent effect by an electron-withdrawing reaction centre cannot be described by equation (4). In fact the data available for electron-withdrawing substituents give no evidence for such a decrease.

On the other hand, for some of the electron-donating substituents Ehrenson *et al.*[3] give σ_R^- values that are less negative than their corresponding σ_R^0 values. The less negative values can be explained by what Taft once called,[2b] ' a " levelling effect " on the substituent electron-release by a powerfully electron-releasing reaction centre '.

This might be interpreted as a reversal of the classical direct resonance effect. If this is possible for electron-donating substituents, there is no reason why it should be impossible for electron-withdrawing substituents.

The arguments given above show that, although the formulation of the discussed extensions of the Hammett relation is relatively simple, the actual state of the description of substituent effects is still complicated. The form of equation (4) indicates that a description by one single equation of all possible kinds of substituent effects leads to an extension of the Taft equation of the general form (8). Such an equation should be used with

$$\log (K_X/K_H) = \sum_{i=1}^{i=n} \rho_i\sigma_i \quad (8)$$

as few restrictions as possible, while n (the number of substituent parameters) must be small enough to make the description useful for practical purposes, and large enough to give a reasonably precise description of the experimental data.

The aim of this work is to find optimal values of substituent constants that can be used in such a general equation.

Another problem encountered in the application of the Hammett equation or its extensions is outlined below. In the literature of the last two decades one can find different numerical values for each of the substituent parameters mentioned above (*e.g.* compare the σ_I values of Taft *et al.*[2b] with those of Ehrenson *et al.*[3]). In general, such differences can be caused by incorporation of new theoretical concepts or new experimental data in the derivation of σ values.

Some examples are given below. Charton [10] derived σ_I from pK values for the dissociation of substituted acetic acids. Originally, Taft did not use these data for his set of σ_I values, because he expected that in this equilibrium steric effects would be present. Though both sets of σ_I values show good agreement, the differences may still be caused by minor steric effects.

The original Hammett σ_m and σ_p values were for the greater part derived directly from pK values for the

---

* Equation (4) cannot be used in regression analysis. handling ρ and r as adaptable parameters, because the numerical values of, *e.g.*, σ_R^0 and $\Delta\sigma_R^+$, are strongly correlated (see refs. 3 and 4*b*).

dissociation of substituted benzoic acids in water as measured by Dippy *et al.*[11]. A number of authors have presented pK values for the same equilibrium, with a precision at least comparable to the data of Dippy, *e.g.*, from the data of Briegleb and Bieber,[12,*] and of Bolton *et al.*,[13] σ_m values for the nitro group can be derived of 0.763 and 0.742, respectively, in rather poor agreement with Hammett's value of 0.710. Even more obvious are the differences in the proposed values of σ_p for the iodine substituent, which range from 0.276 (given by Hammett and supported by measurements of Robinson and Ang [14]) to 0.18.[5c,15]

To eliminate the effect of experimental errors, Jaffé [5a] suggested that a substituent constant should be redefined as ' the value of σ which best fits the entire body of the experimental data '. McDaniel and Brown [15] followed this suggestion with regard to the σ_m and σ_p values, but they still used Dippy's data as a standard for a limited number of substituents. For $\sigma_R{}^0$ parameters an optimalization procedure, based upon data for several reaction series, was performed by Ehrenson.[16] Ehrenson *et al.*[3] reported optimalization procedures for all the parameters of the Taft equation. However, in these procedures several arbitrary restrictions were applied, *e.g.*, $\sigma_R(\text{BA}) \equiv \sigma_R{}^-$ for the substituents CH_3, F, Cl, and Br. Recently, Sjöström and Wold [17] presented optimized values for $\sigma_m{}^0$ and $\sigma_p{}^0$ which have restricted validity. In this work we adhere to the proposition of Jaffé, formulated as follows: substituent constants, to be used in a general equation of the type (8), must be based upon a large number of experimental data, and not upon data for a few typical reaction series only. To ensure the general applicability of such an equation and such constants, no arbitrary restrictions must be imposed on the derivation procedure. These conditions are the basis of the statistical procedure presented in this work.

Selection of Literature Data.—The data selected for the statistical procedure were chosen on the basis of the following considerations.

(a) Only data for chemical equilibria or reaction rates were used. Numerous other data can be described by Hammett type substituent constants, *e.g.*, n.m.r. shifts,[3] i.r. spectral data,[3] ionization potentials,[18] *etc.* However, it remains mostly unclear why rules, that hold for chemical equilibria or reactions, should hold also for these physical quantities. Moreover, such data can partly represent substituent effects that are specific for the physical quantity in question, *e.g.* the frequently mentioned magnetic anisotropy effects in n.m.r. spectroscopy.[19] Therefore it was thought prudent to exclude these data from the generation of substituent constants; afterwards it can be tested whether the constants are applicable to data that were excluded from the procedure.

(b) The data set contains for the greater part data for *meta-* and *para*-monosubstituted benzene derivatives. As to disubstituted derivatives: though sometimes it is found that Hammett type relations are well applicable, there are reasons to exclude these data. In general, it must be assumed that, despite the incidental validity of additivity rules,[20] the effect of a substituent may be perturbed by a second substituent. *E.g.*, an adjoining substituent can influence the resonance effect of the considered substituent, *viz.*, the sterical hindrance of resonance of the nitro group by the methyl group in 3,5-dimethyl-4-nitrobenzoic acid.[21] This hindrance is absent if the nitro group is replaced by a halogen atom. Because such interactions are specific for a particular combination of substituents, general rules are not expected to hold.[22]

(c) All types of reactions, as defined by Ehrenson *et al.*,[3] are represented in the procedure to ensure the general applicability of the generated substituent constants. Any doubt as to the correctness of classification of a particular reaction series will not affect the precision of the final descriptions with the substituent constants. For the number of reaction series of each type, see Table 4. The *meta*-substituted reaction series were considered to be of the same type as the corresponding *para*-substituted series. (In the statistical procedure the data for *meta-* and *para*-substituted reactions were treated as independent series.) The two σ_I type reaction series † are the dissociation of 4-X-bicyclo-[2.2.2]octane-1-carboxylic acids [25] (series 1) and the dissociation of 4-X-quinuclidinium ions [26] (series 2). Further, in all groups of reaction types, those series are included that are generally considered to be reference reactions for the type in question, *e.g.*, the dissociation of phenylacetic acids [27] and the alkaline hydrolysis of phenylacetic ethyl esters [4b] ($\sigma_R{}^0$), the dissociation of benzoic acids in water [11] [σ_R(BA)], the solvolysis of t-cumyl chlorides [9b] and the dissociation of pyridinium ions [28] ($\sigma_R{}^+$), and the dissociation of anilinium ions [8,29] and phenols [30] ($\sigma_R{}^-$).

(d) For statistical reasons, a reaction series must contain a sufficient number of data (substituents). No strict rules were applied, and for reasons of diversity [following from point (c)] a number of series were included, for which only four suitable data were available. A close watch was kept on the results of the statistical procedure, to verify that the final description of these data gave no unlikely results: from this description estimates of unavailable data can be computed, and these should match either with estimates from other descriptions, *e.g.*, the Hammett relation, or else with other literature data [see point (e)].

(e) As a rule, series of data that might be composed of results of several independent investigations were excluded from the selection to avoid systematic errors, caused by differences in the experimental procedure. Additional data were used for the check on the procedure, mentioned in point (d).

---

[*] These data were judged to be highly reliable by G. Kortüm, W. Vogel, and K. Andrussow, *Pure Appl. Chem.*, 1966, **1**, 187.

† Inclusion of aliphatic series in our procedure implies the assumption that in aliphatic and aromatic series one common type of substituent effect is operational. This assumption is doubted by, *e.g.* Sjöström and Wold,[23] but defended by Topsom.[24] This controversial point will be discussed later.

(f) The data must be sufficiently reliable. Here, too, no strict rules were applied. In general, series containing a great number of data were preferred to short series of greater precision.

(g) A series must contain data for both mesomeric electron-withdrawing and -donating substituents, to be sure that all types of substituent effects are represented. (The variety and number of substituents, proposed by Ehrenson *et al.*[3] for their correlations, is not practical. Only very few series would satisfy both his proposition and the rules mentioned here.) Only two series used in the calculations do not meet this condition, and in this case the results of the statistical procedure were checked as mentioned in point (d).

(h) The number of selected series, *i.e.* 76, must be restricted to limit the required computer time. In general, there are much more series available in the literature that satisfy the rules given in points (a)—(g). However, the inclusion of, *e.g.*, additional series of data for the dissociation of benzoic acids would scarcely give any extra information because there is generally excellent correlation between series of data for benzoic acids. Our choice of σ_R^0 type series may contain strongly correlated information. This had to be so, because substituent effects in these series are relatively small: therefore a large number of data must be included in the procedure to give these data an acceptable statistical weight.

(i) The number of substituents, for which constants can be computed (*i.e.* 17), is restricted by the consideration given above [points (a), (b), (d)—(f)]. Those substituents were chosen, for which at least ten data, acceptably distributed over the reaction types, were available.

Statistical Procedure and Results.—The literature data are of the type $Y_{r,s}$, where r and s denote the reaction and the substituent under consideration, respectively, and Y represents the substituent effect log (K_X/K_H).* Equation (8) is reformulated here as (9). R and S will

$$Y_{r,s} = \sum_{i=1}^{n} R_{r,i}S_{i,s} \qquad (9)$$

be treated as adjustable parameters, depending on the reaction and the substituent, respectively.

If data were available for all combinations of r and s, optimal values of R and S might be found by a variant of factor analysis.[31,†] The number of terms n on the right hand side of equation (9) is then chosen so large as appears necessary to reduce the residuals [equation (10)] to an acceptably small level. However, data are

$$Y_{r,s} - \sum_{i=1}^{n} R_{r,i}S_{i,s} = \mathrm{res}_{r,s(n)} \qquad (10)$$

missing for more than 50% of all combinations of r and s. Therefore, factor analysis cannot be applied directly and a special procedure was devised, which has been described elsewhere.[32] (Three parameters, $S'_{1,s}$, $S'_{2,s}$, and $S'_{3,s}$, were obtained by the programs MISFAC–NEWFAC described in ref. 32, and were used to complete the data matrix. The factor analysis applied to the completed data matrix actually rotates the parameters $S'_{1,s}$, $S'_{2,s}$, and $S'_{3,s}$ to the principal components $S_{1,s}$, $S_{2,s}$, and $S_{3,s}$.)

The sums of squares $(\sum_{r,s} R_{r,i}\, S_{i,s})^2$, explained by successive factors R_iS_i in equation (9), are given in Table 1.

TABLE 1

Sums of squares explained by successive factors R_iS_i (all series used), and $R_i''S_i''$ (only aromatic series used) in equation (9). The sums of squares have been found by application of factor analysis

i	$\sum_{r,s} (R_{r,i}S_{i,s})^2$	$\sum_{r,s} (R_{r,i}''S_{i,s}'')^2$
1	1 234.58	1 246.09
2	283.66	260.50
3	19.75	18.51
4	0.32	0.22
5	0.22	0.17
6—17	0.54	0.46
$\sum_{r,s} Y^2_{r,s}$	1 539.08	1 525.95

It is clear from the results in Table 1 that factors higher than three do not give a substantial improvement of the description of the data: the fourth factor explains only 30% of the residual sum of squares, unexplained by the first three factors (*i.e.*, 1.08). Moreover, only three series contribute 0.24 to the total sum of squares explained by the fourth factor, *i.e.* 0.32, which means that the fourth factor improves mainly the description of these three series. From the chemical point of view, these series have not much in common.

Therefore it can be concluded that a fourth factor, as occurs in equation (4), is not necessary to describe substituent effects in general.

To check whether exclusion of the aliphatic series seriously alters the results of the procedure, we repeated it, now leaving out the data on the two aliphatic series. The resulting parameters will be denoted as S_1'', S_2'', and S_3''. It is clear from the results in Table 1, third column, that again three parameters should be used to describe the data. The sum of squares of the residuals, $\Sigma\mathrm{res}^2_{r,s(3)}$ [see equation (10)], for the 74 aromatic series is 0.97 for S_1, S_2, and S_3, and 0.85 for S_1'', S_2'', and S_3'', a slightly better result for the latter parameters. To decide whether the aliphatic series 1 and 2 ‡ should be included or excluded from the procedure, we calculated the standard deviations of the regressions of the data for series 1 and 2 on S_1, S_2, and S_3, and on S_1'', S_2'', and S_3''. [For comparison we also calculated how well Taft's σ_I and σ_R^0 (see Table 3) can be described by these two sets of parameters.] The results are given in Table 2.

For series 1 it is clear that S_1'', S_2'', and S_3'' give even

---

* All data for the substituents CF_3 to NO_2 (see Table 3) were made positive to facilitate comparisons between data series. So, some data series are actually values of log (K_H/K_X).

† In the variant of factor analysis applied in this work, the data matrix is used unmodified for obtaining a covariance matrix in the initial step of the analysis. *i.e.* the rows or columns in the data matrix were not standardized.

‡ See ' Selection of literature data,' point (c).

TABLE 2

Standard deviations of regressions on S_1, S_2, and S_3, and on S_1'', S_2'', and S_3''

	sd (S_1, S_2, S_3)	sd (S_1'', S_2'', S_3'')
Series 1 [a,b]	0.066	0.054
Series 2 [a,c]	0.086	0.256
Series 2 minus C_6H_5, NH_2	0.064	0.095
σ_I (Taft)	0.029	0.040
σ_I (Taft) minus C_6H_5, NH_2	0.027	0.028
$\sigma_R{}^0$ (Taft)	0.046	0.061
$\sigma_R{}^0$ (Taft) minus C_6H_5, NH_2	0.034	0.035

[a] See ' Selection of literature data,' point (c). [b] Substituents included CH_3. C_2H_5, Cl, Br, OCH_3. OH, CF_3, and NO_2. [c] Substituents included all substituents of Table 3 except F and CF_3.

a slightly better description than S_1, S_2, and S_3. For series 2 the description by S_1'', S_2'', and S_3'' is inferior to that by S_1, S_2, and S_3. Inspection of the data showed that, using S_1'', S_2'', and S_3'', the largest residual occurred for NH_2, and the second largest for C_6H_5. This is illustrated by the results of the regressions in Table 2 where the data for these substituents have been omitted. The same situation occurs with σ_I and $\sigma_R{}^0$.

TABLE 3

Substituent constants found in this work and corresponding literature values

X	S_1	S_2	S_3	σ_I	σ_I (Taft) [3]	σ_R	$\sigma_R{}^0$ (Taft) [3]	σ_E
CH_3	−0.0955	−0.0822	0.0074	0	−0.04	−0.113	−0.11	−0.040
C_2H_5	−0.0920	−0.0805	0.0194	0		−0.107		−0.050
Pr^i	−0.0885	−0.0723	0.0394	−0.009		−0.093		−0.067
Bu^t	−0.0913	−0.0536	−0.0376	−0.019		−0.100		0.012
C_6H_5	0.0060	−0.1564	0.1427	0.153	0.10	−0.086	−0.11	−0.175
F	0.1099	−0.2609	−0.4534	0.463	0.50	−0.286	−0.34	0.384
Cl	0.1720	−0.2218	−0.2896	0.450	0.46	−0.177	−0.23	0.243
Br	0.1823	−0.2163	−0.2675	0.450	0.44	−0.162	−0.19	0.225
I	0.1748	−0.2010	−0.1940	0.414	0.39	−0.132	−0.16	0.157
OCH_3	−0.1577	−0.3756	−0.0839	0.279	0.27	−0.143	−0.45	−0.033
OH	−0.2192	−0.4293	−0.0410	0.274		−0.475		−0.095
NH_2	−0.4037	−0.5572	0.2907	0.188	0.12	−0.580	−0.48	−0.468
CF_3	0.3346	−0.1101	−0.2666	0.474	0.45	0.003	0.08	0.270
$COCH_3$	0.2252	−0.1329	0.4887	0.271	0.28	0.139	0.16	−0.475
CO_2CH_3	0.2588	−0.0779	0.2480	0.282	0.30	0.134	0.14	−0.226
CN	0.4143	−0.1836	0.1037	0.566	0.56	0.085	0.13	−0.095
NO_2	0.4726	−0.2104	0.3197	0.613	0.65	0.152	0.15	−0.302

TABLE 4

Number of series within a group of related series having a significant ρ_E at the given confidence level

Series number	Reaction type	Substituent position	Number of series in group	Confidence levels (%)		
				90	95	99
1—2	σ_I	4	2	0	0	0
2—31	$\sigma_R{}^0$	3	13	1	0	0
		4	16	6	5	1
32—46	σ_R (BA)	3	8	3	3	1
		4	7	3	3	2
47—60	$\sigma_R{}^+$	3	6	6	5	2
		4	8	8	8	8
61—76	$\sigma_R{}^-$	3	7	3	0	0
		4	9	8	8	6

Because there is no reason to assume the aliphatic literature data on NH_2 and C_6H_5 to be erroneous, we think it justified to include all data in series 1 and 2 in the generation procedure. This yields parameters S_1, S_2, and S_3 that give a good description of aliphatic substituent effects of NH_2 and C_6H_5. The description of aromatic data is hardly influenced by the inclusion of the aliphatic data.

The set of S_1, S_2, and S_3 values, resulting from the factor analysis, is presented in Table 3. These parameters have no physical, but only statistical meaning. In order to find parameters that correspond to existing substituent parameters (*i.e.*, Taft's σ_I and $\sigma_R{}^0$ values), the rotation procedure described in the Appendix was applied to the set of S values.

The set of σ_I, σ_R, and σ_E values, resulting from the rotation procedure, is given in Table 3, columns 5, 7, and 9.* The regression equation to be used with these constants is (11). With the three σ parameters of Table

$$\log (K_X/K_H) = \rho_I\sigma_I + \rho_R\sigma_R + \rho_E\sigma_E \qquad (11)$$

3 and the regression equation (11), ρ values and standard deviations of the regression for all series in the data matrix were calculated. Then we checked for which series the term $\rho_E\sigma_E$ gives a significant contribution to the description of the experimental data. According to Snedecor and Cochran [33] this can be done by applying student's *t* test to ρ_E/sd(ρ_E), where sd(ρ_E) is the standard deviation of the regression coefficient ρ_E. The results are presented in Table 4.

DISCUSSION

For the aliphatic series 1 and 2 the ρ_E (and ρ_R) values do not differ significantly from zero. Therefore, the precision of the description does not change if instead of a three-parameter regression a one-parameter regression of the data on σ_I is applied. For the *meta*-substituted $\sigma_R{}^0$ type series (3—31), the ρ_E values do not differ signi-

---

* The same standard deviation of the regression is found if in equation (11) the set of S_1, S_2. and S_3 values is used instead of the set of σ_I, σ_R, and σ_E values.

ficantly from zero, and thus the precision of the description is not significantly impaired if a two-parameter regression of the data on σ_I and σ_R is applied, instead of a three-parameter regression. The same conclusion does not hold for the group of *para*-substituted σ_R^0 series, for the group of σ_R(BA) type reaction series (32—46) or for the entire group of all *meta*-substituted reaction series (see Table 4).

The ρ_I values for sets of corresponding *meta*- and *para*-substituted reaction series are consistently of the same magnitude. An important observation is that all *para*-substituted σ_R^- type reactions have negative ρ_E values, while all σ_R^+ type reactions (series 47—60) have positive ρ_E values.* Thus, the sign of ρ_E can be used to identify a particular reaction with respect to reaction type. Further relevant discussion is hardly possible, especially for σ_R^+ and σ_R^- type reactions. As stated in the Appendix, a rotation leading to a parameter that represents the pure direct resonance effect can not be performed at the present time. Therefore, the σ_E values are a mixture of unknown composition of the inductive effect, the σ_R^0 type resonance effect and the direct resonance effect. This means that any search for a superficial chemical plausibility in the presented list of σ_E parameters is futile, and that for the present, chemical interpretations of ρ values must be tentative if ρ_E is of a significant magnitude.

TABLE 5

Experimental and calculated values for log K_X^D/K_H^D of *ortho*-substituted pyridines in polyethylene glycol

	165.5 °C			206 °C		
Substituents	Exp.	Calc. [equation (11)]	Calc.[b] [equation (2)]	Exp.	Calc. [equation (11)	Calc.[b] [equation (2)]
CH_3	−0.155	−0.177	−0.128	−0.138	−0.152	−0.051
C_2H_5	−0.194	−0.180	−0.128	−0.167	−0.154	−0.051
F	−0.033	−0.007	−0.171	0.002	0.016	−0.037
Cl	0.010	−0.007	−0.079	0.018	0.015	−0.004
Br	0.026	−0.005	−0.054	0.033	0.016	0.004
OCH_3	−0.567	−0.552	−0.433		−0.460	−0.149
CN	−0.025	−0.017	0.177	0.007	0.012	0.098
sd [a]		0.029	0.139		0.017	0.090
ρ_I		−0.052	0.203		0.002	0.133
		±0.040 [c]	±0.166 [c]		±0.024	±0.114 [c]
ρ_R		1.229	0.478		1.052	0.181
		±0.066	±0.129		±0.084	±0.136
ρ_E		0.962			0.823	
		±0.076			±0.072	

[a] Standard deviations of the regression of the experimental data on the substituent constants involved in the particular regression equation. [b] Regression on σ_I and σ_R^+ values, obtained from ref. 3. As no σ values are given for C_2H_5 we used in our correlations for C_2H_5 the σ_I and σ_R^+ values for CH_3. Exclusion of the experimental data on C_2H_5 from all correlations resulted in: sd [equation (11). 165.5 °C] = 0.030, sd [equation (2), 165.5 °C] = 0.152; sd [equation (11). 206 °C] = 0.016. sd [equation (2), 206 °C] = 0.073. The corresponding calculated values for log $K_{OCH_3}^D/K_H^D$ at 206 °C were: −0.418 [equation (11)] and −0.075 [equation (2)]. [c] Standard deviations of the regression coefficients ρ.

To illustrate the usefulness of the substituent constants derived in this work, we determined data on the substituent effect on gas–liquid phase equilibria for substituted phenols, anilines and pyridines. (Details are given in ref. 34.) These data are related by equation (13)

$$K_X^D/K_H^D = V_N^{Q,X}V_N^{Ref,H}/V_N^{Q,H}V_N^{Ref,X} \qquad (13)$$

where K_X^D/K_H^D is the substituent effect on the distribution constant of the phase equilibrium, V_N is the g.l.c. net retention volume for a particular stationary phase and temperature, Q,H and Q,X are the unsubstituted and substituted investigated substances, respectively, and Ref,H and Ref,X are unsubstituted and substituted reference substances, respectively. Phenols, anilines, and pyridines were the investigated substances; the correspondingly substituted benzenes were used as the reference substances.

For the non-protic stationary phase polyethylene glycol, and experimental temperatures of 165.5 and 206 °C, we found that log K_X^D/K_H^D values for *meta*- and *para*-substituted anilines, phenols, and pyridines could be described well, by equation (11), but also by the Taft equation (2). The best fits by the Taft equation were obtained with σ_R(BA)[3] for *meta*-substituted series and σ_R^-(A)[3] for *para*-substituted series.

The standard deviations of the three parameter regressions were of the same magnitude as the experimental error, and turned out to be slightly but significantly smaller than the standard deviations of the best fits of the Taft equation (*F* test, 95% confidence level).

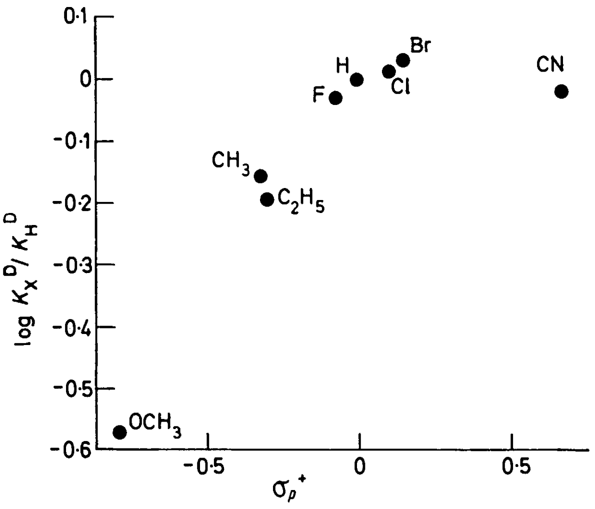
The data on *ortho*-substituted pyridines at 165.5 and 206 °C are given in Table 5; those at 165.5 °C are also shown in the Figure as a function of σ_p^+. These data could be described acceptably well by equation (11). (See Table 5; compare sd with an estimated experimental error of ±0.012.) Substituting the obtained ρ values and the tabulated σ values for OCH_3 in equation (11), we calculated a value of log $K_{OCH_3}^D/K_H^D$ at 206 °C of −0.460 (see Table 5). This calculated value is in excellent agreement with an estimation of −0.462, obtained from an extrapolation to 206 °C of data measured at 130 and 165.5 °C. (This extrapolation is based upon the theory of g.l.c. and requires no knowledge about substituent effects.) The best fit to the experimental data on *ortho*-substituted pyridines by the Taft equation

---

* This is not a consequence of any restriction imposed on the generation of the substituent or reaction constants: in the rotation procedure no assumptions regarding series 32—76 were used. It must be recalled that the substituent effect was defined in such a way that all ρ_I and ρ_R values became positive.

(2) was obtained with σ_R^+ parameters;[3] the results are also given in Table 4. It is evident that, with regard to the data for *ortho*-substituted pyridines, the results of the three parameter regression are superior to the results obtained by applying the Taft equation. The Figure shows that the Hammett equation also does not give an adequate description of the data. To put it more strongly: we are convinced that no other proposed equation is able to give a description of the data on *ortho*-substituted pyridines, and a prediction of log $K_{OCH_3}{}^D/K_H{}^D$, as precise as equation (11) does.

The success of the application of equation (11) to the data shows that the system contains only the common kinds of substituent effects, *i.e.*, inductive, resonance, and direct resonance effects, be it in a mixture that is not often encountered in other systems.



Hammett plot of log $K_X{}^D/K_H{}^D$ for *ortho*-substituted pyridines in polyethylene glycol at 165.5 °C *versus* σ_p^+ (ref. 5c)

Conclusions.—(1) From 570 literature data on substituent effects for 76 reactions and 17 substituents optimal values of the parameters ρ and σ for the description of the data by equation (8) can be generated by a statistical procedure that includes factor analysis.

(2) Three substituent constants σ are required to describe substituent effects in all known types of reactions and equilibria. For σ_I type reactions one constant, σ_I, comparable to Taft's σ_I, is sufficient. For *meta*-substituted σ_R^0 type reactions two constants, σ_I and σ_R, comparable to Taft's σ_I and σ_R^0, are required. For *para*-substituted σ_R^0 type reactions, and for σ_R(BA), σ_R, and σ_R type reactions three constants, σ_I, σ_R, and σ_E, are necessary. The substituent constant σ_E represents a mixture of unknown composition of the inductive effect, the σ_R^0 type resonance effect, and the direct resonance effect. The sign of the ρ_E reaction constant is indicative of the reaction type: it is negative for *para*-substituted σ_R^- reactions and positive for σ_R^+ reactions.

APPENDIX

Rotation procedure for the conversion of the set of S_1, S_2, and S_3 values to a set of σ_I, σ_R, and σ_E values.

Phase 1.—In terms of matrix algebra, we have a matrix Y of data $Y_{r,s}$, a matrix S of substituent parameters $S_{i,s}$, a matrix R of reaction parameters $R_{r,i}$ † and a matrix D of residuals $\mathrm{res}_{r,s(i)}$. They are related by expression (14).

$$Y - R \,.\, S = D \qquad (14)$$

The matrix D meets the condition that, for $i = 3$ equation (15) applies. As equation (16) holds where T is an orthogonal 3 × 3 rotation matrix and T' is its transpose, it is clear that not only the obtained matrices R and S, but also each matrix $R^* = R \,.\, T$ and the corresponding matrix $S^* = T' \,.\, S$, yield the matrix D *via* equation (14).

$$\sum_{r,s} \mathrm{res}_{r,s(i)}{}^2 = \text{minimal} \qquad (15)$$

$$R \,.\, S = R \,.\, T \,.\, T' \,.\, S \qquad (16)$$

If we accept Taft's hypothesis that the data in series 1—31 (see Table 4) can be described by two factors, the first aim of the rotation procedure is to find a matrix R^* of rotated reaction parameters $R^*_{r,i}$ that are about zero for $r =$ 1—31 and $i = 3$. This can be accomplished by factor analysis of the matrix $R_{31.3}$, that is composed of the first 31 rows of the matrix R. As a result, the matrices T' and $R^*_{31.3}$ are found.

Phase 2.—As the elements of the third column of the matrix $R^*_{31.3}$ are very small, approximation (17) holds

$$R^*_{31.3} \,.\, S^*_{3.17} \simeq R^*_{31.2} \,.\, S^*_{2.17} \qquad (17)$$

where $R^*_{31.2}$ is the matrix that is obtained by deleting the third column of $R^*_{31.3}$, and $S^*_{2.17}$ is obtained by deleting the last row of $S^*_{3.17}$. Analogous to equation (16), (18) holds

$$R^*_{31.2} \,.\, S^*_{2.17} = R^*_{31.2} \,.\, U_2 \,.\, U'_2 \,.\, S^*_{2.17} \qquad (18)$$

where U_2 is an orthogonal 2 × 2 rotation matrix and U'_2 is its transpose. Thus, not only the matrices $R^*_{31.2}$ and $S^*_{2.17}$, but also each matrix $R^{**}_{31.2} = R^*_{31.2} \cdot U_2$ and the corresponding matrix $S^{**}_{2.17} = U_2 \cdot S^*_{2.17}$ meet our requirements. If we accept the hypothesis that the data in series 1 and 2 can be described by one factor (*i.e.*, an inductive parameter, see Table 3), the second aim of the rotation procedure is to find a matrix R^{**} of rotated reaction parameters $R^{**}_{r,i}$ that are about zero for $r =$ 1—2 and $i = 2$.

This can be accomplished by factor analysis of the matrix $R^*_{2.2}$, which is composed of the first two rows of the matrix $R^*_{31.2}$. As a result, the matrices U'_2 and $R^{**}_{2.2}$ are found.

The results of the orthogonal rotation procedure up to this point can be summarized by equations (19) and (20).

$$R^{**} = R^* \cdot U = R \cdot T \cdot U \qquad (19)$$

In the next two phases of the rotation procedure the requirements of orthogonality and normality are partly abandoned.

Phase 3.—In order to have a closer conformity between the obtained inductive substituent parameters $S^{**}_{1,s}$ and Taft's σ_I values, the scale of the $S^{**}_{1,s}$ values was adjusted. The $S^{**}_{1,s}$ values multiplied by the regression coefficient a from the regression of σ_I on $S^{**}_{1,s}$ (13 σ_I values of Ehrenson *et al.*[3]) [equation (21)]. These values of $a \,.\, S^{**}_{1,s}$ will be called from here onwards σ_I values, because of their close resemblance to Taft's σ_I values (see Table 2, columns 5 and 6).

Phase 4.—In a similar way a parameter comparable to

---

† Obtained by multiple regression analysis of Y_s on $S_{i,s}$.

$$\begin{pmatrix} R^{**}_{1.1} & \delta & \delta \\ R^{**}_{2.1} & \delta & \delta \\ R^{**}_{3.1} & R^{**}_{3.2} & \delta \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ R^{**}_{31.1} & R^{**}_{31.2} & \delta \\ R^{**}_{32.1} & R^{**}_{32.2} & R^{**}_{32.3} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ R^{**}_{76.1} & R^{**}_{76.2} & R^{**}_{76.3} \end{pmatrix} = \begin{pmatrix} R^{*}_{1.1} & R^{*}_{1.2} & \delta \\ R^{*}_{2.1} & R^{*}_{2.2} & \delta \\ R^{*}_{3,1} & R^{*}_{3.2} & \delta \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ R^{*}_{31.1} & R^{*}_{31,2} & \delta \\ R^{*}_{32.1} & R^{*}_{32.2} & R^{*}_{32.3} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ R^{*}_{76.1} & R^{*}_{76.2} & R^{*}_{76.3} \end{pmatrix} \cdot \begin{pmatrix} U_{1.1} & U_{1.2} & 0 \\ U_{2.1} & U_{2.2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (20)$$

$$S^{**} = U' \,.\, T' \,.\, S$$

Taft's $\sigma_R{}^0$ was found from a regression of $\sigma_R{}^0$ on $S^{**}{}_{1.s}$ and $S^{**}{}_{2.s}$ (13 $\sigma_R{}^0$ values of Ehrenson *et al.*[3]) [equation (22)]. The values of $b \,.\, S^{**}{}_{1.s} + c \,.\, S^{**}{}_{2.s}$ will be called in the following σ_R values.

$$\sigma_I = a \,.\, S^{**}{}_{1.s} \qquad (21)$$

$$\sigma_R{}^0 = b \,.\, S^{**}{}_{1.s} + c \,.\, S^{**}{}_{2.s} \qquad (22)$$

The values of $S^{**}{}_{3.s}$, from here onwards referred to as σ_E values (E from ' extra '), were kept unaltered (*i.e.*, the σ_E values are for the 17 substituents mathematically orthogonal to the σ_I and σ_R values) because for the third substituent effect no satisfying datum can be found in the literature.† Thus, the σ_E values represent a mixture of unknown composition of the inductive effect, the $\sigma_R{}^0$ type resonance effect and the effect of direct resonance.

In terms of matrix algebra, the operations performed in phases 3 and 4 of the rotation procedure lead to the calculation of a matrix S^{***} of σ_I, σ_R, and σ_E values from equation (23). As a consequence the corresponding matrix R^{***} of ρ_I, ρ_R, and ρ_E values is given by equation (24).

$$S^{***} = V \,.\, S^{**} = \begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ c & 0 & 1 \end{pmatrix} \,.\, S^{**} \qquad (23)$$

$$R^{***} = R^{**} \,.\, V^{-1} = R^{**} \,.\, \begin{pmatrix} a^{-1} & 0 & 0 \\ -a^{-1} \,.\, b \,.\, c^{-1} & c^{-1} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (24)$$

Supplementary material is given in Supplementary Publication No. SUP 22452 (14 pp.).‡

[8/153 *Received*, 30*th January*, 1978]

---

† The idea of combining $S^{**}{}_{1.s}$, $S^{**}{}_{2.s}$, and $S^{**}{}_{3.s}$ values in a parameter that resembles an existing direct resonance parameter, *e.g.*, $\Delta\sigma_R{}^+$ of Yukawa and Tsuno, was rejected for the reason given in the note to equation (4).

‡ For details of Supplementary Publications see Notice to Authors No. 7 in *J.C.S. Perkin II*, 1978, Index issue.

REFERENCES

[1] L. P. Hammett, ' Physical Organic Chemistry,' McGraw-Hill, New York, 1940, p. 184.

[2] (*a*) R. W. Taft and I. C. Lewis, *J. Amer. Chem. Soc.*, 1958, **80**, 2436; (*b*) R. W. Taft, C. N. Deno, and P. S. Skell, *Ann. Rev. Phys. Chem.*, 1958, **9**, 287.

[3] S. Ehrenson, R. T. C. Brownlee, and R. W. Taft, *Progr. Phys. Org. Chem.*, 1973, **10**, 1.

[4] (*a*) Y. Yukawa and Y. Tsuno, *Bull. Chem. Soc. Japan*, 1959, **32**, 971; Y. Yukawa, Y. Tsuno, and M. Sawada, *ibid.*, (*b*) 1966, **39**, 2274; (*c*) 1972, **45**, 1198.

[5] (*a*) H. H. Jaffé, *Chem. Rev.*, 1953, **53**, 191; (*b*) P. R. Wells, *ibid.*, 1963, **63**, 171; (*c*) C. D. Ritchie and W. F. Sager, *Progr. Phys. Org. Chem.*, 1964, **2**, 323; (*d*) O. Exner, 'Advances in Linear Free Energy Relationships,' eds. N.B. Chapman and J. Shorter, Plenum, New York, 1972, p. 1; (*e*) M. Charton, *Chem. Tech.*, 1974, **4**, 502: (*f*) E. Grunwald and J. E. Leffler, in ' Physical Chemistry,' ed. H. Eyring, Academic Press, New York, 1975. vol. 7, p. 171.

[6] (*a*) C. G. Swain and E. C. Lupton, *J. Amer. Chem. Soc.*, 1968, **90**, 4328; (*b*) M. J. S. Dewar, R. Golden, and J. M. Harris, *ibid.*, 1971, **93**. 4187.

[7] A. J. Hoefnagel and B. M. Wepster, *J. Amer. Chem. Soc.*, 1973, **95**, 5357.

[8] A. Bryson, *J. Amer. Chem. Soc.*, 1960, **82**, 4858.

[9] (*a*) H. C. Brown and Y. Okamoto, *J. Amer. Chem. Soc.*, 1957, **79**, 1913; (*b*) H. C. Brown, Y. Okamoto, and T. Inukai, *ibid.*, 1958, **80**, 4964; (*c*) Y. Okamoto, T. Inukai, and H. C. Brown, *ibid.*, p. 4969.

[10] M. Charton, *J. Org. Chem.*, 1964, **29**, 1222.

[11] (*a*) J. F. J. Dippy and F. R. Williams, *J. Chem. Soc.*, 1934, 1888; (*b*) J. F. J. Dippy, F. R. Williams, and R. H. Lewis, *ibid.*, 1935, 343; (*c*) J. F. J. Dippy and R. H. Lewis, *ibid.*, 1936, 644; (*d*) J. W. Baker, J. F. J. Dippy, and J. E. Page, *ibid.*, 1937, 1774; (*e*) L. G. Bray, J. F. J. Dippy, and S. R. C. Hughes, *ibid.*, 1957, 265; (*f*) J. F. J. Dippy, S. R. C. Hughes, and B. C. Kitchiner, *ibid.*, 1964, 1275.

[12] G. Briegleb and A. Bieber, *Z. Elektrochem.*, 1951, **55**, 250.

[13] P. D. Bolton, K. A. Fleming, and F. M. Hall, *J. Amer. Chem. Soc.*, 1972, **94**, 1033.

[14] R. A. Robinson and K. P. Ang, *J. Chem. Soc.*, 1959, 2314.

[15] D. H. McDaniel and H. C. Brown, *J. Org. Chem.*, 1958, **23**, 420.

[16] S. Ehrenson, *Tetrahedron Letters*, 1964, 351.

[17] M. Sjöström and S. Wold. *Chemica Scripta*, 1976, **9**, 200.

[18] A. Streitwieser, *Progr. Phys. Org. Chem.*, 1963, **1**, 1.

[19] (*a*) P. Diehl, *Helv. Chim. Acta*, 1961, **44**, 829; (*b*) R. F. C. Brown, L. Radom, S. Sternhell, and I. D. Rae, *Canad. J. Chem.*, 1968, **46**, 2577.

[20] P. D. Bolton and F. M. Hall, *J. Chem. Soc.* (*B*), 1970, 1247.

[21] J. P. Schaefer and T. J. Miraglia, *J. Amer. Chem. Soc.*, 1964, **86**, 64.

[22] (*a*) R. W. Taft and H. D. Evans, *J. Chem. Phys.*, 1957, **27**, 1427; (*b*) M. J. S. Dewar and Y. Takeuchi, *J. Amer. Chem. Soc.*, 1967, **89**, 390.

[23] M. Sjöström and S. Wold, *Chemica Scripta*, 1974, **6**, 114.

[24] R. D. Topsom, *Progr. Phys. Org. Chem.*, 1976, **12**, 1.

[25] (*a*) H. D. Holtz and L. M. Stock, *J. Amer. Chem. Soc.*, 1964, **86**, 5188; (*b*) F. W. Baker, R. C. Parish, and L. M. Stock, *ibid.*, 1967, **89**, 5676.

[26] A. Ceppi, W. Eckhardt, and C. A. Grob, *Tetrahedron Letters*, 1973, 3627.

[27] (*a*) J. F. J. Dippy and F. R. Williams, *J. Chem. Soc.*, 1934, 161; (*b*) J. F. J. Dippy and R. H. Lewis, *ibid.*, 1937, 1008; (*c*) A. J. Hoefnagel, J. C. Monshouwer, E. C. G. Snorn, and B. M. Wepster, *J. Amer. Chem. Soc.*, 1973, **95**, 5350.

[28] A. Fischer, W. J. Galloway, and J. Vaughan, *J. Chem. Soc.*, 1964, 3591.

[29] A. de Courville, *Compt. rend.*, 1966, **262**, 1196.

[30] C. van Hooidonk and L. Ginjaar, *Rec. Trav. chim.*, 1967, **86**, 449.

[31] H. H. Harman, ' Modern Factor Analysis,' Univ. Chicago Press, Chicago, 1970.

[32] C. L. de Ligny, N. G. van der Veen, and J. C. van Houwelingen, *Ind. Eng. Chem. Fund.*, 1976, **15**, 336.

[33] G. W. Snedecor and W. G. Cochran, ' Statistical Methods,' Iowa State College Press, 1956, 5th edn., pp. 418—419.

[34] (*a*) G. H. E. Nieuwdorp, C. L. de Ligny, and N. G. van der Veen, *J. Chromatography*, 1978, **154**, 133; (*b*) G. H. E. Nieuwdorp, Thesis, 1976, University of Utrecht, The Netherlands.