L. Ethyl 2-Carboethoxy-5-fluoro-3-indoleacetate.—A mixture of 4-fluoroaniline,[30] 67 ml. of concd. hydrochloric acid and 100 ml. of water was diazotized in the usual manner with 14 g. of sodium nitrite in 25 ml. of water. The condensation and cyclization process was accomplished by the King and L'Ecuyer[31] procedure. The diazonium solution was added slowly to a mixture of 48 g. of ethyl α-acetoglutarate,[32] 200 ml. of ethanol and 150 ml. of 20% sodium hydroxide at 0–5° with stirring. Thirty minutes after the addition, the mixture was acidified with hydrochloric acid and a dark red oil separated which partially solidified on standing. The precipitate was dissolved in ether, dried and the ether was evaporated. The dried residue was dissolved in 100 ml. of absolute ethanol, dry hydrogen chloride gas was passed in to saturation, and the mixture was then refluxed for one hour. Upon pouring into ice, a brown semi-solid separated which upon recrystallization from aqueous ethanol gave yellow needles. Vacuum sublimation gave pure ethyl 2-carboethoxy-5-fluoro-3-indoleacetate as white needles, m.p. 123–124°, yield 15 g. (30%).

M. 2-Nitro-5-fluorophenylpyruvic Acid.—The procedure of Meyer and Balle[33] was adapted to the synthesis of this compound. To 150 m . of absolute ethanol, 14 g. of clean sodium chips was added slowly and allowed to react completely. With stirring and cooling, 88 g. of diethyl oxalate was added slowly and followed by a solution of 78 g. of 2-nitro-5-fluorotoluene[34] in 150 ml. of anhydrous ether. No evidence of sodium enolate precipitation was observed, and the ether was evaporated to reduce the volume. The concentrate was acidified by pouring into a ice-hydrochloric acid mixture, whereupon a red oil separated. The oil was collected in ether, and two extractions with 200-ml. portions of sodium hydroxide (*N*) solution removed the alkali-soluble material. Acidification of the alkaline extract precipitated the crude pyruvic acid. Recrystallization from benzene gave the pure compound as white needles, m.p. 146–147°, yield 38 g. (31%).

N. 5-Fluoro-2-indolecarboxylic Acid.—The preceding pyruvic acid was converted to an indole derivative by the Cornforth and Robinson[35] procedure. Ten grams of 2-nitro-5-fluorophenylpyruvic acid was added to 2.2 g. of sodium hydroxide in 85 ml. of water. With stirring, 30 g. of sodium hydrosulfite dihydrate was added slowly and the reaction was slightly exothermic. Stirring was continued until a test sample gave no red color in excess alkali solution. The mixture was acidified with hydrochloric acid, heated on a steam-bath to expel sulfur dioxide, cooled, and extracted with ether. The yield of crude product from ether evaporation was 8 g. Recrystallization from aqueous ethanol followed by vacuum sublimation (140° (1 mm.)) gave white granular crystals as pure 5-fluoro-2-indolecarboxylic acid. The compound does not melt, but at 245° it appears to decarboxylate to the fluoroindole.

O. Bis-(3-fluorophenyl)-phosphinic acid was prepared by the general procedure described by Doak and Freedman.[36] To 14 g. of phosphorus trichloride and 29 g. of cuprous bromide in 200 ml. of dry ethyl acetate, 42 g. of 3-fluorobenzenediazonium fluoroborate[24] was added gradually at 50°. The reaction mixture was refluxed two hours and steam distilled. The residue was evaporated to about 50 ml. and, on cooling, the diarylphosphinic acid crystallized. Recrystallization from low boiling petroleum ether with a few drops of benzene gave white needles, m.p. 166–167°.

Tabulation of Experimental Data.—The tables contain information on the physical constants and notes on the experimental work on the compounds studied. Known compounds are listed in cases where they were required to operate a given synthesis or where the literature data are inadequate.

The syntheses notes indicate the parent compound required in the synthesis, the procedure or general method of preparation, the recrystallization solvents for final purification, and the yield data based on the parent compound. In some cases molar ratios of reactants and reaction temperatures are given. The tables are interrelated through compound numbers, thus making it possible to trace the ancestry of a given compound back to a known starting material. The previous information can be illustrated as follows: compound no. 21 (2-fluoro-4,5-dichlorophenoxyacetic acid) was prepared from parent compound no. 63 (2-fluoro-4,5-dichlorophenol) by procedure Ja, recrystallized from benzene (footnote *a*), yield 89%. Likewise, the ancestry of parent compound no. 63 can be traced back through its intermediates, 91 and 128, to a known starting material, compound no. 120 (3,4-dichlorofluorobenzene).

---

[30] F. Swarts, *Bull. classe sci. Acad. roy. Belg.*, 241 (1913).

[31] F. E. King and P. L'Ecuyer, *J. Chem. Soc.*, 1901 (1934).

[32] A. H. Blatt, Ed., "Organic Syntheses," John Wiley and Sons, Inc., New York, 1943, Coll. Vol. II, general procedure p. 262.

[33] F. Meyer and G. Balle, *Ann.*, **403**, 188 (1914).

[34] G. Schiemann, *Ber.*, **62B**, 1794 (1929).

[35] R. H. Cornforth and R. Robinson, *J. Chem. Soc.*, 680 (1942).

[36] G. O. Doak and L. D. Freedman, THIS JOURNAL, **73**, 5658 (1951).

URBANA, ILL.

[CONTRIBUTION FROM THE DEPARTMENT OF CHEMISTRY, ILLINOIS INSTITUTE OF TECHNOLOGY]

Multiple Variation in Structure–Reactivity Correlations[1]

BY SIDNEY I. MILLER

RECEIVED DECEMBER 26, 1957

As a natural extension of linear correlations of structure with reactivity, it is proposed that the four-parameter equation $y = px + qxz + rz + s$ be used for dual variations where these are characterized by the structural constants x and z. The theoretical implications of this equation and its capacity to store, predict and evaluate data is indicated. For the rapid review of data in which two structural variations separately give linear correlations, an approximate but convenient graphical method, the "generating procedure," is developed.

Quantitative correlations of chemical data have often been cast in the form of two- or three-parameter equations.[2–4] What are the consequences of introducing multiple structural variations in systems to which these equations apply?

The discussion begins with the straight line

$$y = Ax + B \qquad (1)$$

Consider a specific example in which the Hammett[3] and Swain–Scott[5] equations apply.

$$\text{Hammett} \qquad \log k = \rho\sigma + \log k_0 \qquad (1a)$$

$$\text{Swain-Scott} \qquad \log k = sn + \log k_0 \qquad (1b)$$

The rate constants, k, for a series of reactions

$$R_iC_6H_4E + N_j \longrightarrow \text{Products} \qquad (2)$$

are to be correlated. R_i, the substituent of an

---

[1] Supported by the Office of Ordnance Research, U. S. Army.

[2] L. P. Hammett, "Physical Organic Chemistry," McGraw–Hill Book Co., Inc., New York, N. Y., 1940, Chap. 7; J. Hine, "Physical Organic Chemistry," McGraw–Hill Book Co., Inc., New York, N. Y., 1956, Chaps. 2, 6, 8, 12.

[3] H. H. Jaffé, *Chem. Revs.*, **53**, 191 (1953).

[4] R. W. Taft, Jr., in "Steric Effects in Organic Chemistry," edited by M. S. Newman, John Wiley and Sons, Inc., New York, N. Y., 1956, Chap. 13.

[5] C. G. Swain and C. B. Scott, THIS JOURNAL, **75**, 141 (1953).

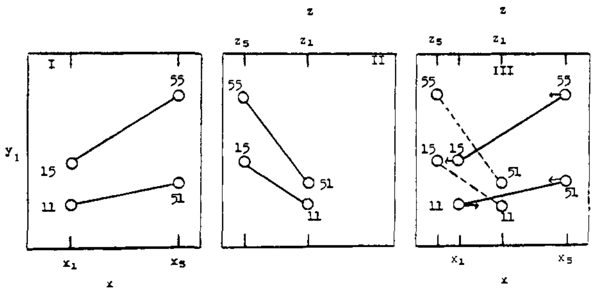


Fig. 1.—Generating procedure: y is linear in x at constant z_1 or z_5 or in z at constant x_1 or x_5. In III the first pair of lines generate the second pair.

aromatic electrophile, is characterized by σ; N, the nucleophile, is characterized by n. It is evident that if i is the number of substituents R_i and if j is the number of nucleophiles N_j, then the total number of rate constants to be correlated is their product ij. Moreover, if 1a and 1b do indeed apply, then j Hammett lines or i Swain–Scott lines are required to store the rate data.

It will be shown that all of these data can be derived when four of the rate constants are known; alternatively, a single four-parameter equation can store all of the data.

Generating Procedure.—The mathematical and theoretical aspects of multiple variation will be discussed later. This section will develop a graphical approach to the interrelation of correlations such as 1a and 1b.

$R_iM + R_jN$ (or $R_iM - NR_j$) represents a generalized system whose properties y_{ij} are being studied. The system undergoes structural variations R_i and R_j characterized by the structural constants x_i and z_j, respectively.[6] *It is postulated that the properties y_{ij} will give linear correlations with x_i or with z_j; either the set of lines contained in equation 3 or the set of lines contained in equation 4 can represent all y_{ij}.* (The subscripts i and j are associated with x and z, respectively, and run from 1 to i or j.) For example, the combinations of R_iM with R_1N yield 3a and the combinations of R_iM with R_5N yield 3b

$$y_{ij} = a_jx_i + b_j \qquad (z_j = \text{constant}) \qquad (3)$$

$$y_{ij} = c_iz_j + d_i \qquad (x_i = \text{constant}) \qquad (4)$$

$$y_{i1} = a_1x_i + b_1 \qquad (z_1 = \text{constant}) \qquad (3a)$$

$$y_{i5} = a_5x_i + b_5 \qquad (z_5 = \text{constant}) \qquad (3b)$$

If all x_i and z_j and 3a and 3b are given, any y_{ij} can be found. To show this, we note that points *11* (x_1, z_1, y_{11}) and *51* (x_5, z_1, y_{51}) define equation 3a; *15* (x_1, z_5, y_{15}) and *55* (x_5, z_5, y_{55}) define equation 3b. But *11* and *15* also define 4a, and *51* and *55* also define 4b.

$$y_{1j} = c_1z_j + d_1 \qquad (x_1 = \text{constant}) \qquad (4a)$$

$$y_{5j} = c_5z_j + d_5 \qquad (x_5 = \text{constant}) \qquad (4b)$$

Pairs of points such as *91* and *95* can be obtained from 3a and 3b; these in turn define the line

$$y_{9j} = c_9z_j + d_9 \qquad (x_9 = \text{constant}) \qquad (4c)$$

In this way all of the i lines of set 4 can be obtained.

The first steps of the graphical generating procedure are shown in Fig. 1.

This procedure can be extremely useful for the rapid scanning of data, the prediction of new data and for the quick test of possible interrelations. For example, equations 1c and 1d

McGary–Okamoto–Brown $\log p_f = \rho\sigma^+$ (1c)
Brown–Nelson $\log p_f = cS_f$ (1d)

should cover most electrophilic substitution in monosubstituted benzenes.[7,8] p_f and m_f are partial rate factors in the p- and m-positions of a monosubstituted benzene. With two lines of the 1c set and the given[7,9] σ^+ and S_f ($S_f = \log p_f/m_f$) values, we generated two lines of the 1d set and from these all other lines of the 1c set. The agreement between predicted and experimental data was excellent. Thus, the generating procedure predicted new data quickly and showed the interrelation between 1c and 1d to be successful. A more critical approach to this system will be made later.

It must be emphasized that these graphical methods are not intended to replace the statistical analysis of data. The generating procedure can be used, however, to help decide whether a statistical analysis is worth making.

In practice, the accuracy of predicted data depends entirely on the accuracy of the four generating points and on the validity of the component linear relations; therefore, whenever possible, the two best least squares lines should be used to generate other data. A mean deviation of ±15%, say, in the parent lines should be reflected in the generated lines; but, in fact, the reliability of the derived data will vary somewhat because the structural constants themselves are subject to uncertainty.

Figures 2–5 show how this procedure was applied to illustrative cases. The generating lines, always number *1* and *2*, or the generating points, always solid circles, are taken from the literature. The predicted data are given in the form of broken lines. These should be compared with observed least squares lines bearing the same number primed or with experimental points.

Turning specifically to the problem of structure-reactivity, there are cases in which two sets of Hammett equations may be interrelated. Rate data for reaction 5 are given in Fig. 2.[10]

$$R_iC_6H_4CONHOCOC_6H_4R_j + NH_3 \longrightarrow R_jC_6H_4COO^- + R_iC_6H_4NCO \qquad (5)$$

In these as well as in later examples, the Hammett σ-values and the "observed" Hammett plots were usually taken from Jaffé's excellent review.[3] It is seen that there is agreement between predicted and experimental data. (This system also has been analyzed statistically according to equation 14c with a correlation coefficient of 0.993.[3])

The generating procedure shows a similar successful interrelation in the system[11]

$$R_iC_6H_4COCl + R_jC_6H_4NH_2 \longrightarrow R_iC_6H_4CONHC_6H_4R_j + HCl \qquad (6)$$

---

(6) The term "structural constant" is more general than "substituent constant." The word "parameter" is more properly associated with A and B of equation 1; in some applications A has been called a "reaction constant," *e.g.*, ρ.

(7) C. W. McGary, Jr., Y. Okamoto and H. C. Brown, THIS JOURNAL, **77**, 3037 (1955).

(8) H. C. Brown and K. L. Nelson, *ibid.*, **75**, 6292 (1953).

(9) H. C. Brown and C. R. Smoot, *ibid.*, **78**, 6255 (1956).

(10) W. B. Renfrow, Jr., and C. R. Hauser, *ibid.*, **59**, 2308 (1937).

(11) F. J. Stubbs and C. Hinshelwood, *J. Chem. Soc.*, S71 (1949); E. G. Williams and C. N. Hinshelwood, *ibid.*, 1079 (1934); Jaffé's log k_0 should be increased by 1.778.[3]

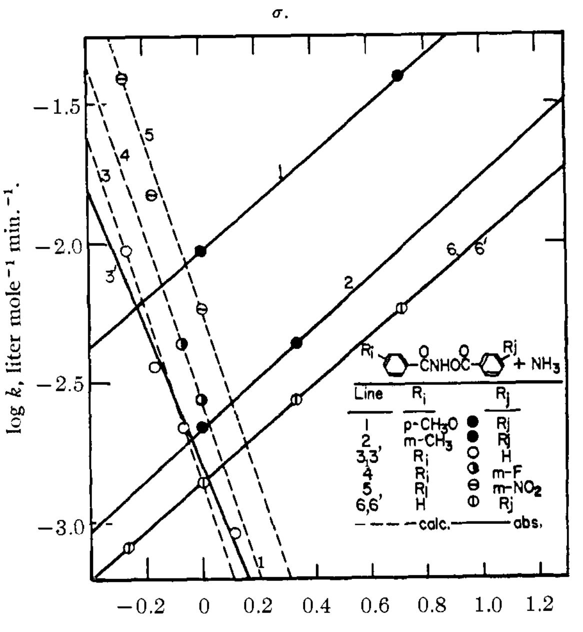


Fig. 2.—Generating procedure: the Lossen rearrangement at 30° (Renfrow, Hauser).

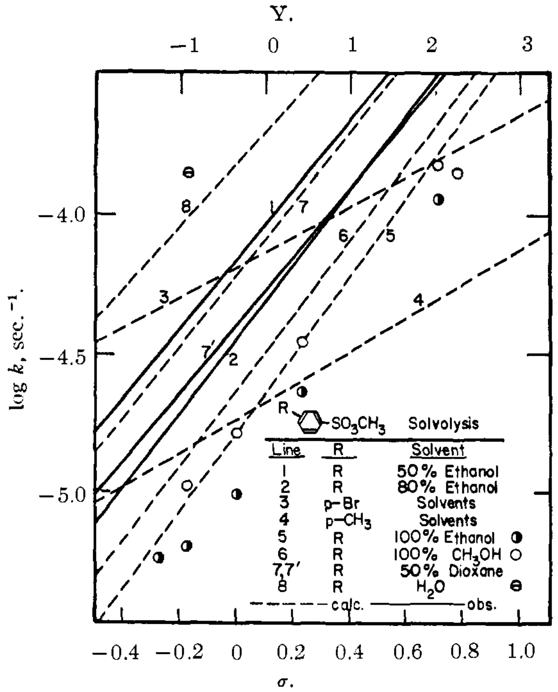


Fig. 3.—Generating procedure: solvolysis of methyl arylsulfonates in various solvents at 50° (Robertson).

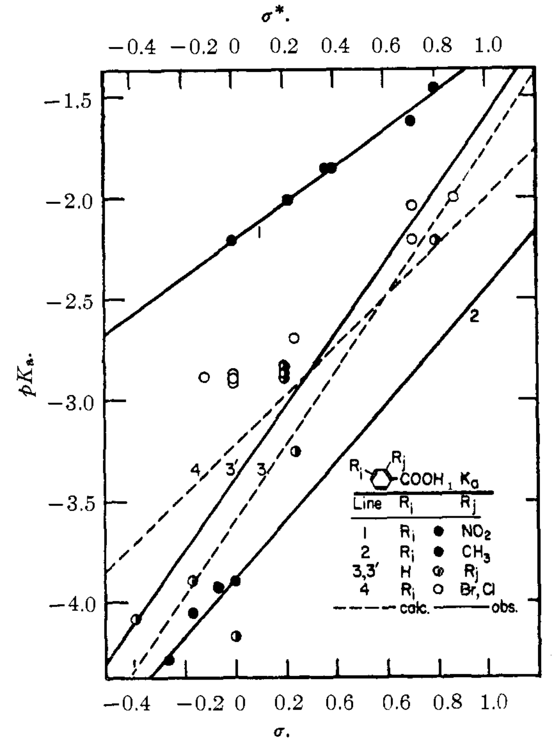


Fig. 4.—Generating procedure: ionization constants of 2,4- or 2,5-benzoic acids in water at 25°.

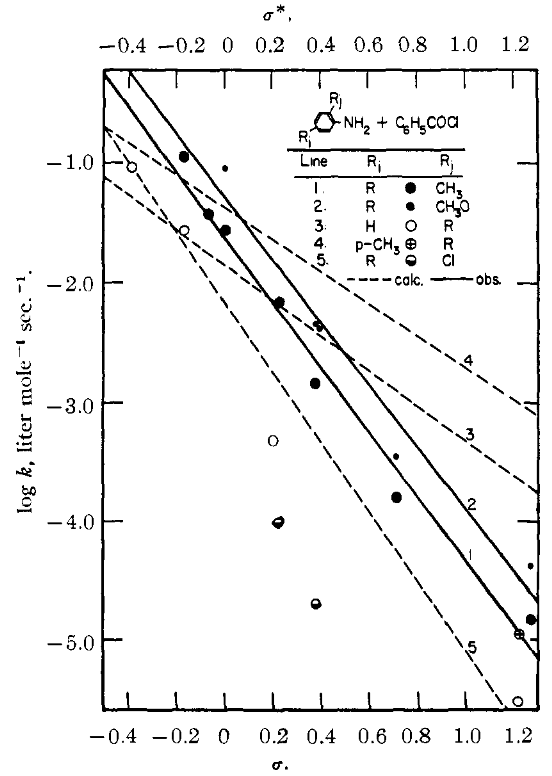


Fig. 5.—Generating procedure: benzoylation of 2,4- or 2,5-anilines in benzene at 25° (Hinshelwood, Stubbs, Williams).

This correlation will be discussed further in a later section.

In Fig. 3, the Hammett and the Grunwald–Winstein equations are interrelated.[12]

Grunwald-Winstein $\log k = mY + \log k_0$ (1e)

The reaction is the solvolysis of $R_iC_6H_4SO_3CH_3$ in mixed solvents at 50°. Solvolytic constants Y were taken from Grunwald and Winstein[12] and the kinetic data were taken from Robertson.[3,13] The range in k is a factor of 100. Since it has recently

---

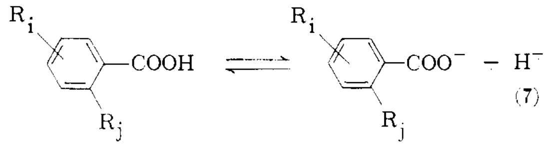
(12) E. Grunwald and S. Winstein, THIS JOURNAL, **70**, 846 (1948).

(13) R. E. Robertson, *Can. J. Chem.*, **31**, 589 (1953); **33**, 1536 (1955).

been suggested that 1e correlates data best for various mixtures of a single solvent in water,[14] it is gratifying that the predicted data lie within a factor of 1.4–1.5 in k. It is possible that a wider range of solvents would reveal greater deviations.

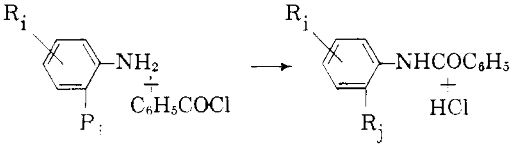
In the next two figures, the Hammett and one Taft equation are interrelated.[4]

Taft $\qquad \log k = p^*\sigma^* + \log k_0 \qquad (1f)$



(7)

The σ_j^* are taken from Taft.[4] In Fig. 4 equilibrium data for the reaction at 25° in water are shown. Most of the K_a were taken from Shorter and Stubbs[3,15]; a few were taken from Dippy's review.[16] Since Taft does not list a σ^* for *o*-hydroxy, data for such acids were not considered. A range of 1000 in K_a is covered. It turns out that line *3* predicted for *o*-benzoic acids falls within the precision limits, a factor of 1.4 in K_a, set by Taft in his calculation of line *3'*.[4] While the data for line *4* do not scatter as badly, the predicted line *4* is off by a factor of *ca.* 1.6 in K_a. Data for the reaction



(8)

are given in Fig. 5.[3,11] A 40,000-fold factor in k is covered. Taft was able to correlate the data for the reaction of four *o*-substituted anilines with benzoyl chloride to within a factor of 2 in the rate constants. His least squares line for the data (open circles) is not shown but would have been close to line *5*. All of the predicted data deviate appreciably from the experimental data. It would appear that the Taft polar energy equation does not correlate the effects of *o*-substituents on k in this example and possibly on K_a of the preceding example.

Mathematics of Multiple Structural Variation

The Generating Procedure, Analytically.—The assumptions have been given in the preceding section in italics. Generalizing the subscripts 1 and 5 to *m* and *n*, the previously chosen generating points *11*, *51*, *15* and *55* become *mm*, *nm*, *mn* and *nn*, respectively. At the appropriate values of the structural constants x_m and x_n, equations 3a and 3b, can be solved for y_{mm}, y_{nm}, y_{mn} and y_{nn}. These points permit the transformations

$$c_m = (y_{mm} - y_{mn})/(z_m - z_n) \text{ and}$$
$$d_m = (z_m y_{mn} - z_n y_{mm})/(z_m - z_n) \qquad (9)$$

Since all c and d parameters can be obtained in this way, all of the lines of the second set (equation 4) can be given.

Although there is considerable choice as to the generating points, there are important restrictions. Indeed, once the experimenter is aware of these, he can design his research so as to get the maximum amount of information. Two points on a line and any two other points not on the line can generate all of the lines. If, however, no two of the four points are collinear, then graphical methods cannot be used: the four points define a surface, equation 13, from which all lines can be generated algebraically. Further, if two generating lines are to be used, they cannot be chosen from different sets. Lines 3a and 4b, for example, are defined by three points *mm*, *nm* and *nn* lying in the plane

$$y_{ij(i = n \text{ or } j = m)} = a_m x_i + c_n z_j + (b_m - c_n z_m) \qquad (10)$$

Since experimental data often include at least one standard substituent for which $x_1 = z_1 = 0$, equation 10 often takes the form

$$y_{ij(i \text{ or } j = 1)} = a_1 x_i + c_1 z_i + y_{11} \qquad (10a)$$

It is of interest that when lines of the parent set are parallel with slope a, then all lines of the other set are parallel with slope c

$$c = (b_m - b_n)/(z_m - z_n) \text{ and}$$
$$a = (d_m - d_n)/(x_m - x_n) \qquad (11)$$

For parallel sets with the same structural constants ($x_i = z_i$), the symmetric points y_{11}, y_{22} --- y_{nn} lie on the line

$$y_{ii} = (a + c)x_i + y_{11} \qquad (x_1 = 0) \qquad (12)$$

Four-parameter Equation.—The fact that 3a and 3b generate all y_{ij} suggests that one four-parameter equation should do this as well, but more efficiently. Such is

$$y = px + qxz + rz + s \qquad (13)$$

a hyperbolic paraboloid in the family of quadric surfaces.

The problem of simultaneous variation in two structural constants has been treated previously.[3,4] The equation used was of the type

$$y = px + rz + s \qquad (14)$$

At this point two non-equivalent alternatives should be distinguished: either 13 or 14 may be *postulated* for the correlation of data. As it is usually intended that the complex equation reduce to a set of linear equations of the type 1 and that the same set of structural constants apply throughout, 13 must be used. The distinction between 10 and 14 appears not to have been generally recognized. Because previous treatments have been inadequate, equation 13 should perhaps be justified.[3–5,14]

The property y depends on two independent factors. Thus

$$y = f(x,z) \quad \text{and} \quad dy = \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial y}{\partial z}\right)_x dz \qquad (15)$$

The general solution of 15 should include all equations represented by 3 and 4. Such a solution is equation 13. Since

$$\left(\frac{\partial y}{\partial x}\right)_z = p + qz \quad \text{and} \quad \left(\frac{\partial y}{\partial z}\right)_x = r + qx \qquad (16)$$

one can recover two sets of equations

$$y_{im} = (p + qz_m)x_i + \text{constant and}$$
$$y_{nj} = (r + qx_n)z_j + \text{constant} \qquad (17)$$

which are equivalent to 3 and 4. Equation 14 is a particular solution of 15; since $q = 0$, one recovers from 16 the sets of parallel lines

---

[14] S. Winstein, A. H. Fainberg and E. Grunwald, THIS JOURNAL, **79**, 4146 (1957).

[15] J. Shorter and F. J. Stubbs, *J. Chem. Soc.*, 1180 (1949).

[16] J. F. Dippy, *Chem. Revs.*, **25**, 151 (1939).

$$y_{\rm lm} = px_{\rm l} + \text{constant} \quad \text{and} \quad y_{\rm nj} = rz_{\rm j} + \text{constant} \qquad (18)$$

Though any four non-coplanar points can define the surface 13, the points previously used in the generating procedure were used to reveal the composition of the parameters.

$$p = \frac{a_{\rm m}z_{\rm n} - a_{\rm n}z_{\rm m}}{z_{\rm n} - z_{\rm m}} \qquad q = \frac{a_{\rm n} - a_{\rm m}}{z_{\rm n} - z_{\rm m}}$$
$$r = \frac{b_{\rm n} - b_{\rm m}}{z_{\rm n} - z_{\rm m}} \qquad s = \frac{b_{\rm m}z_{\rm n} - b_{\rm n}z_{\rm m}}{z_{\rm n} - z_{\rm m}} = y_{11} \qquad (19)$$

There is one other interesting manipulation of 13. If this surface is moved to new axes (as in Fig. 6), one obtains

$$y' = qx'z' \qquad (20)$$

The same sort of equation results when the intercepts of the parent correlations are zero as in 1c and 1d

$$\log p_{\rm f} = q\sigma^{+}S_{\rm f} \qquad (21)$$

If it is desired to vary the structure but keep the property y constant, Fig. 6 should be useful. The allowable variations are immediately evident. By the same token, if z must be kept constant, the choice of y and x becomes very much restricted.

Higher-parameter Equations.—Equations 13, 14 and 22 correlate a set of properties y to several structural constants[17–19]

Swain–Mosely–Brown $\log k = c_1d_1 + c_2d_2 + \log k_0$ (14a)

Edwards $\log k = \alpha E_{\rm n} + \beta H + \log k_0$ (14b)

Hansson $\log k = r\,(s_{\rm a} + s_0) + t + \log k_0$ (22)

The arguments given previously can be extended to the interrelations of 13 with 1. This amounts to a simultaneous threefold variation in structure

$$y = ax + bz + cw + dxz + exw + fzw + gxzw + h \qquad (23)$$

x, z and w are structural constants and the parameters a-h differ from those in 13 and 1. Admittedly, examples for which 23 is suitable are scarce, but the appearance of 22 and the proliferation of 14 in various forms suggest that 23 may become useful. The surfaces, 13 and 23, having few parameters, nevertheless represent many systems or actually families of systems. Though their cross-terms make for complexity, they are the simplest equations which express multiple structural variation and yet reduce rigorously to straight lines.

Concluding Remarks.—The generating procedure is neither as elegant nor as precise as equation 13. It is subject to the errors inherent in the generating points or lines and in the structural constants used. Apart from its pictorial directness, the chief advantage of the graphical generating procedure lies in its speed and simplicity. On the other hand, 13 represents the best fit for all of the available data. Once obtained, it becomes a compact and convenient method for the storage, prediction and evaluation of data.[20]

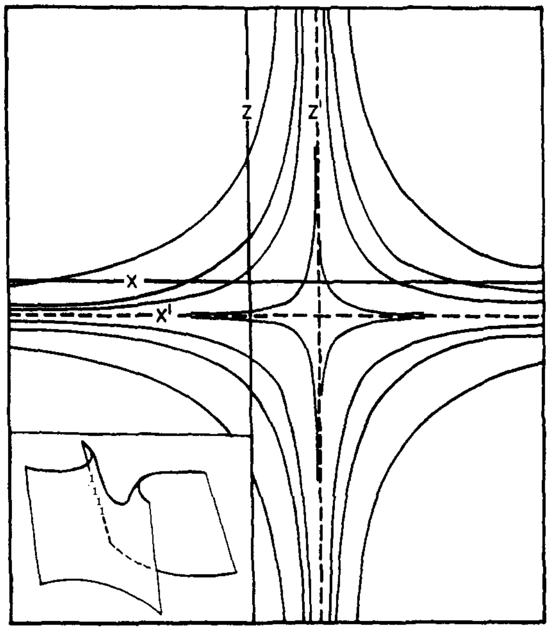


Fig. 6.—Contours at constant y for dual structure variations given by $y = px + qxz + rz + s$, quadrants 1 and 3 or quadrants 2 and 4. Inset figure is the surface given by the four-parameter equation.

Our work is a development of Jaffé's work on multiple structural variation and the Hammett equation. His most general equation was

$$\log k = \rho_1\sigma_1 + \rho_2\sigma_2 + \log k_0 \qquad (14c)$$

Jaffé pointed out that data which fall on intersecting lines, as in equation 12, did not constitute a fair test of equation 14c.[21] He also recognized empirically that 14c should be obeyed if the reaction constants (the $a_{\rm j}$ of our equation 3) did not vary, that is, if the lines of the set were parallel.[22] However, when $q \neq 0$ and when wide structural variations are examined, equation 13 is necessary.

We cite two systems for which $q \neq 0$. Jaffé has given a statistical analysis of the benzoylation rates of anilines, reaction 6, at $R_{\rm i}$ = H or $R_{\rm j}$ = H.[3,11] He has suggested that 14c does not correlate all of the available data satisfactorily. However, using his $\rho_1 = p$ and $\rho_2 = r$, we calculate $q = -0.266$. Since eleven of the twelve rate constants can be correlated by 14c or by 10a, the statistical significance of q is uncertain.

Brown, *et al.*, have treated rate data of several electrophilic reactions of aromatic systems statistically and have listed ρ values appropriate to 1c. When these ρ's are divided by the appropriate $S_{\rm f}$, the parameter q of 21 can be evaluated. The results are

$-\rho^{a\ b}$	4.62	6.53	11.35	4.59	4.30	9.57
$S_{\rm f}^{c}$	1.11	1.24	2.44	0.96	1.36	2.09
$-q$	4.16	5.27	4.65	4.78	3.16	4.58

Mean $q = -4.43 \pm 0.52$

[a] H. C. Brown and Y. Okamoto, THIS JOURNAL, **70**, 1913 (1957). [b] H. C. Brown and F. R. Jensen, *ibid.*, **80**, 2296 (1958). [c] Calculated from partial rate factors for toluene from reference given.[a,b]

---

(17) C. G. Swain, R. B. Mosely and D. E. Brown, THIS JOURNAL, **77**, 3731 (1955).

(18) J. O. Edwards, *ibid.*, **76**, 1540 (1954); **78**, 1819 (1956).

(19) J. Hansson, *Svensk Kem. Tidskr.*, **66**, 351 (1954); **67**, 246 (1955).

(20) The author would appreciate suggestions on old or new systems suitable for processing by equations 13 or 23.

(21) W. A. Pavelich and R. W. Taft, Jr., THIS JOURNAL, **79**, 4935 (1957).

(22) C. K. Hancock and J. S. Westmoreland, *ibid.*, **80**, 545 (1958).

If uncertainties in S_f and σ^+ are ignored, one can predict *para* partial rate factors to within about 12% in this set. Here again the available data are as yet limited; moreover, the σ^+-constants are derived from one of the reactions, the solvolysis of phenyldimethylcarbinyl chlorides and the selectivity factors, S_f, derive from the reactions of *m*- and *p*-methyl reactants.

When a pair of linear correlations does not interrelate well, the validity of one or both is in doubt. In Fig. 5, for example, the Hammett least squares lines *1* and *2* may be taken as adequate. These generate lines *3* and *4*. If the Taft polar energy equation 1f applies and if his *o*-methyl and *o*-methoxy σ^*-values are satisfactory, *3* and *4* should fall close to the experimental data. They do not. Line *5*, which derives from *3* and *4*, also fails. For these reasons it was concluded that the Taft equation did not apply in this case. The foregoing suggests an internal consistency check or a validity test for new linear correlations.

For the field of linear correlations it is significant that in equation 13 two factors affect the property y of a system and become interlocked. This is clear from the expressions in 19. Also, the equation of one set of intercepts has the form

$$y_{1j} = rz_j + s = b_j \quad (x_1 = 0) \qquad (24)$$

That is, the spacings of the intercepts of one set at $x_1 = 0$ are proportional to related structural constants of the second set.[23] Any theoretical progress in the Hammett equation, say, can perhaps be carried over to any other equation which it will interrelate. Furthermore, our understandings of structural properties and line parameters which have heretofore developed independently, should become unified.[3]

There is, of course, nothing in equation 13 which restricts it to structure–reactivity problems. With the proper linear correlation and the generating procedure, we have satisfactorily predicted the boiling points of the esters R_iCOOR_j.[24] In chelate chemistry, linear correlations of pK_s (K_s = stability constant) with pK_a or pK_b of the ligand, with ionization potentials of the metal, with reciprocal metal radii, etc., probably could be treated by the methods of this paper.[25] There is a possible application in gas–liquid chromatography which involves separation factors *versus* other variables.[26] Indeed, whenever a data pattern appears similar to that formed by one set of lines as in Fig. 3, the possibility of a second linear correlation should at least be recognized.

Acknowledgment.—We wish to thank Prof. J. Berman of our Mathematics Department for a helpful discussion.

CHICAGO 16, ILL.

(23) The absence of a continuous change in b_j with z_j reported for the chlorination of four series of compounds of the type p-$R_jC_6H_4OCH_2$-$C_6H_4R_j$ indicates a serious discrepancy in an otherwise excellent study.[3]

(24) Y. Ogata and M. Tsuchida, *Ind. Eng. Chem.*, **49**, 415 (1957).

(25) A. E. Martell and M. Calvin, "Chemistry of the Metal Chelate Compounds," Prentice–Hall, Inc., New York, N.Y., Chaps. 4, 5, 1952; L. G. Van Uitert and W. C. Fernelius, THIS JOURNAL, **76**, 379 (1954).

(26) G. J. Pierotti, C. H. Deal, E. L. Derr and P. E. Porter, *ibid.*, **78**, 2989 (1956); A. B. Littlewood, C. S. G. Phillips and D. T. Price, *J. Chem. Soc.*, 1480 (1955).

[CONTRIBUTION FROM THE RADIATION LABORATORY,[1] UNIVERSITY OF CALIFORNIA, BERKELEY]

The Distribution of Radioactivity in Toluene Formed from Benzene and Photolyzed Diazomethane-C^{14}

BY RICHARD M. LEMMON AND WALTER STROHMEIER[2]

RECEIVED MAY 26, 1958

Diazomethane-C^{14} was photolyzed (both by visible and ultraviolet light) in benzene solution. The resultant, labeled toluene was degraded in order to determine the distribution of radioactivity between the methyl group and the ring. No detectable activity was found in the ring; this result indicates the operation of a different mechanism from that observed in the production of labeled toluene by the interaction of benzene with recoiling C^{14} atoms. In addition, it was found that X-rays have, at most, only a negligible effect in promoting the decomposition of diazomethane in benzene solution.

A number of reports have appeared recently describing the irradiation of organic compounds with recoiling C^{14} atoms, derived from the N^{14} (n,p)C^{14} reaction,[3] or with C^{14} ions in a mass spectrometer.[4] These reports have heightened interest and speculation concerning the kind of interaction that is taking place when the "hot" atom or ion is reduced in energy to the point (probably 25–50 e.v.) where a chemical bond formation can take place.

Wolf and co-workers[5] have shown that when toluene is formed from the irradiation of benzene with recoiling C^{14} atoms, not all of the C^{14} activity is present in the methyl group. About 86% of the activity was found in this position and the remaining 14% in the ring. A similar result was obtained for *o*-xylene formed from C^{14}-atom irradiation of toluene.

Doering and Knox[6] photolyzed diazomethane in the presence of benzene and formed toluene and cycloheptatriene. Wolf has concluded[5] that the recoiling C^{14} atom may behave similarly to Doering and Knox's "carbene" (CH_2:) but that the recoil-

(1) The work described in this paper was sponsored by the U. S. Atomic Energy Commission.

(2) Appointment supported by the International Cooperation Administration under a program administered by the National Academy of Sciences; present address: Physikalisch Chemisches Institut der Universität Würzburg.

(3) A. P. Wolf, C. S. Redvanly and R. C. Anderson, THIS JOURNAL, **79**, 3717 (1957), have thoroughly reviewed this type of experiment.

(4) R. M. Lemmon, F. Mazzetti, F. L. Reynolds and M. Calvin *ibid.*, **78**, 6414 (1956).

(5) A. P. Wolf, B. Gordon and C. S. Redvanly, Abstracts of Papers, 131st Meeting of the American Chemical Society, Miami, Fla., April, 1957, p. 12-O.

(6) W. von E. Doering and L. H. Knox, THIS JOURNAL, **75**, 297 (1953).