

# Ionization of Carbonyl Compounds in Sulfuric Acid. Correction for Medium Effects by Characteristic Vector Analysis

John T. Edward\* and Sin Cheong Wong

*Contribution from the Department of Chemistry, McGill University, Montreal,  
Quebec, Canada H3A 2K6. Received December 6, 1976*

**Abstract:** The ultraviolet spectra of 16 carbonyl compounds (ketones, acids, esters, amides, and one aldehyde) in various concentrations of sulfuric acid have been analyzed by the methods of multivariate analysis to separate the effect of protonation from the generalized "medium effect". The protonation of most of the compounds followed neither  $h_0$  nor  $h_A$  but some intermediate acidity function. Protonation of molecules of very small volume followed an acidity function increasing with acid concentration less rapidly than  $h_A$ ; an explanation based on more extensive solvation of the protonated bases is advanced.

In 1935 Flexser, Hammett, and Dingwall<sup>1</sup> showed how the base strength of a weak Bronsted base such as acetophenone could be determined from the changes in the ultraviolet spectrum of the compound when dissolved in aqueous sulfuric acid of increasing acid concentration. Over a limited range of acid concentrations these changes were large, because of progressive protonation of the carbonyl group. However, superimposed on the large effects of protonation were smaller "medium effects" which prevented the family of spectral curves from passing through an isobestic point. These medium effects were also manifested at lower acid concentrations, in which protonation of the carbonyl group was inappreciable, and at higher acid concentrations, in which protonation was essentially complete, by small shifts of absorption peak position and intensity with changing acid concentration.

The problem of medium effects has remained a vexation (Hammett<sup>2</sup> devotes seven pages to it in the second edition of his classic text), particularly in studies of aromatic ketones, where it is fairly large. Consequently, the question of whether the protonation of these compounds is governed by the  $H_0$ ,  $H_A$ , or some other acidity function has been answered differently by different authors<sup>2-15</sup> and sometimes differently by the same authors at different times.<sup>12,13</sup>

Various methods of compensating for the medium effect have been advanced,<sup>1,12,16-20</sup> but the most powerful method,<sup>21</sup> applied by Reeves<sup>22</sup> to the spectral changes attending the ionization of phenylazobenzenes and phenylazonaphthols in sulfuric acid, has been neglected because of its alleged mathematical complexity.<sup>12</sup> This is the method of analyzing multivariate response data by characteristic vector analysis.<sup>23</sup> The analysis extracts from the data the minimum number of factors which contain essentially all the information in the original data and has been much used in fields such as psychology (e.g., in the statistical analysis of cheese-tasting results<sup>24</sup>) where multiple factors are to be expected. The chemist has generally striven to devise experimental procedures where only one variable<sup>25</sup> is being altered at a time, and so usually has not been under the same necessity to use this technique as have been workers in geology, biology, psychology, etc. However, lately multivariate analysis (or factor analysis) has been applied to a number of problems in chemistry,<sup>26-28</sup> and it can be expected to become increasingly important.

The technique has been very simply described by Simonds,<sup>21</sup> and so only a brief description need be given here. The absorbance  $A$  at  $r$  wavelengths is obtained in  $n$  different concentrations of acid, and the data are arrayed to form an  $n$ -row,  $r$ -column data matrix. Description of multivariate response data in terms of characteristic vectors involves the determination of linear transformations of the data to an intrinsic minimum number of parameters containing all the information

in the original data. If  $p$  independent factors are involved in generating the absorbance curves, the sample responses at each wavelength for a given acid concentration will be given by

$$A_1 = \bar{A}_1 + c_1v_{11} + c_2v_{21} + \dots + c_pv_{p1} \quad (1)$$

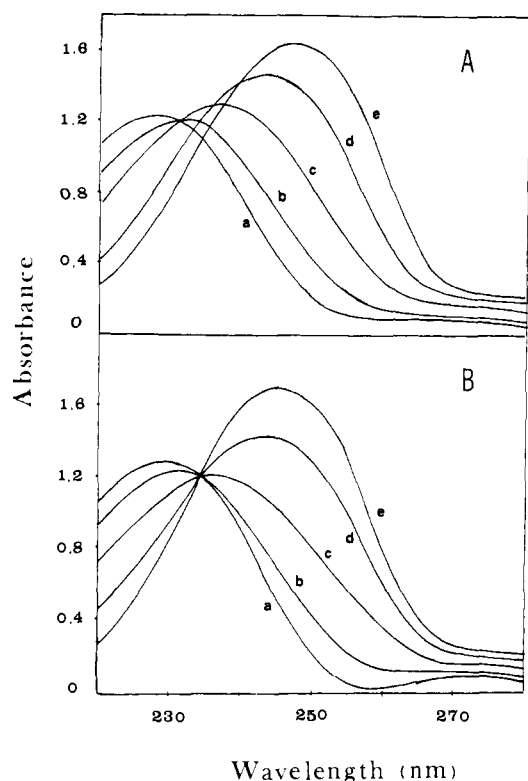
$$A_2 = \bar{A}_2 + c_1v_{12} + c_2v_{22} + \dots + c_pv_{p2} \quad (2)$$

$$A_r = \bar{A}_r + c_1v_{1r} + c_2v_{2r} + \dots + c_pv_{pr} \quad (3)$$

where the choice of  $\bar{A}$ 's is arbitrary (but a set of good  $\bar{A}$  values can greatly reduce the number of mathematical iterations involved in the calculation, and the mean values of the absorbance seem to be a convenient choice). The  $v$ 's are the characteristic vectors, and  $c$ 's are weighting coefficients. Such a set of equations is soluble as long as  $p \leq r$ ; in our case, it is usually found that  $p = 2$ , so that this is possible. A computer program<sup>29</sup> was set up to manipulate the data matrix according to the detailed procedure of Simonds,<sup>21</sup> and it was found that for a large number of carbonyl compounds the absorbance at any acid concentration could be described by a mean absorbance curve  $\bar{A}$  and two characteristic vectors  $v_1$  and  $v_2$ , the first vector accounting for about 96% of the variability with acid concentration and the second about 3% of the variability. It is reasonable on chemical grounds to associate the first vector with the effect of protonation and the second vector with the medium effect, although this is not required by the mathematical analysis.

We can illustrate the method by considering the specific case of benzamide. The experimental absorbance curves in five concentrations of sulfuric acid are given in Figure 1A; curves were also obtained in seven other acid concentrations. The mean curve for the 12 spectra is shown in Figure 2A. The mathematical analysis then shows that any of the experimental curves can be reconstituted by addition to the mean curve of Figure 2A a contribution at each wavelength from the first characteristic vector, given in Figure 2B, weighted by the factor  $c_1$  from Table I, and another contribution at this wavelength from the second characteristic vector, given in Figure 2C, weighted by the factor  $c_2$  from Table I.

Five of the curves reconstituted from the mean curve and the first characteristic vector are shown in Figure 1B; these, like all 12 reconstituted curves, pass through an isobestic point. This gives us confidence that the first characteristic vector is indeed the result of protonation, and that we have to do with a single equilibrium:  $X + H^+ \rightleftharpoons XH^+$ . It should be emphasized that the curves of Figure 1B are obtained by a purely mathematical processing of the raw data by the methods of vector analysis, with no assumptions about the nature of the effects (other than their orthogonality). They do *not* come from an arbitrary shifting of experimental curves to obtain an



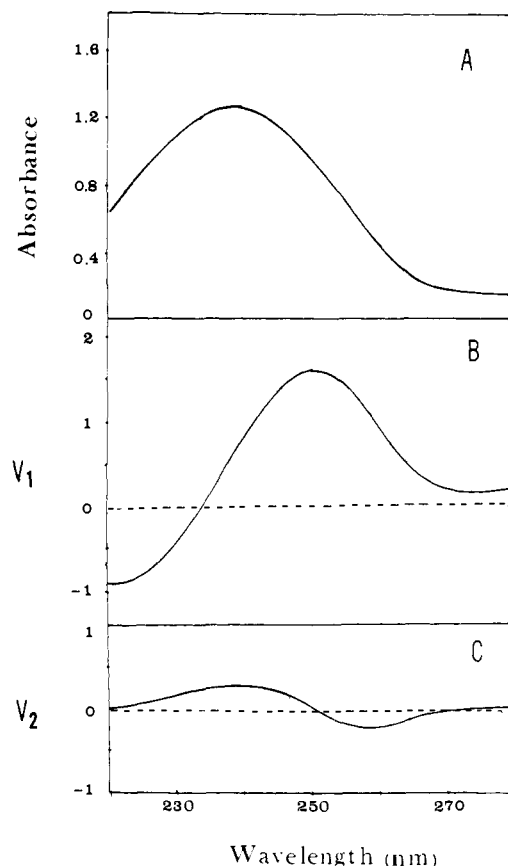
**Figure 1.** Absorption curves of benzamide in aqueous sulfuric acid: (a) 0%; (b) 24.0%; (c) 33.6%; (d) 48.1%; (e) 73.0%. Upper set: data curves. Lower set: curves reconstituted from the mean curve and first characteristic vector.

**Table I.** Weighting Factors  $c_1$  and  $c_2$  of the Characteristic Vectors  $v_1$  and  $v_2$  for the Absorbance of Benzamide in Sulfuric Acid

% H <sub>2</sub> SO <sub>4</sub>	$c_1$	$c_2$
0	-0.457	-0.379
13.0	-0.423	-0.275
24.0	-0.328	-0.051
33.6	-0.191	0.159
36.4	-0.087	0.302
39.7	0.002	0.392
43.0	0.067	0.318
48.1	0.137	0.244
51.7	0.209	0.111
55.3	0.277	-0.015
61.0	0.366	-0.337
73.0	0.427	-0.469

isosbestic point as in the original attempts of Flexser, Hammett, and Dingwall<sup>1</sup> to allow for medium effects. Furthermore, such arbitrary shifting will not work for families of curves crossing over each other not once but two or more times. In such cases the shifts of entire curves required to make the curves pass through an isosbestic point at one wavelength make their deviation from isosbestic points at other wavelengths even worse. This is the case for the spectra of fluorenone and anthraquinone, illustrated in Figures 3A and 4A. On the other hand, application of Simond's method of multivariate analysis yields the families of reconstituted curves shown in Figures 3B and 4B, passing through two and three well-defined isosbestic points.

The spectra of the other carbonyl compounds of Table II were submitted to the same analysis, and all gave families of reconstituted curves which had good isosbestic points. From these reconstituted curves the ionization ratios  $[XH^+]/[X]$  at varying acid concentrations could be obtained. These ion-



**Figure 2.** Mean curve (A) and curves for first characteristic vector  $v_1$  (B) and second characteristic vector  $v_2$  (C) of benzamide.

ization ratios were independent of the wavelength chosen to calculate them, as they should be in the absence of "medium effects". The same ionization ratios could also be obtained directly from the coefficients of the first characteristic vectors without actually reconstituting the spectral curves.

Spectral changes due to causes other than protonation, and loosely defined by the term "medium effects", are represented by the second characteristic vectors. In all cases there is only one "medium effect" vector for each compound (the third and subsequent characteristic vectors are all negligible), indicating that the method of characteristic vector analysis does not recognize the effects of changing medium on the conjugate acid and on the free base as factors independent of each other. They are related and form a single function of acid concentration as indicated by the coefficients of the second characteristic vectors ( $c_2$  in Table I, for instance). We have initiated further work in an attempt to advance our understanding of these "medium effects".

In Figure 5 the values of  $\log [XH^+]/[X]$  of benzaldehyde and five ketones are shown plotted against  $-H_0$ . Equally good linear plots were obtained for the other compounds of Table II. From such plots slopes  $m$  and  $pK_{XH^+}$  values were obtained fitting the equation of Yates and McClelland:<sup>20</sup>

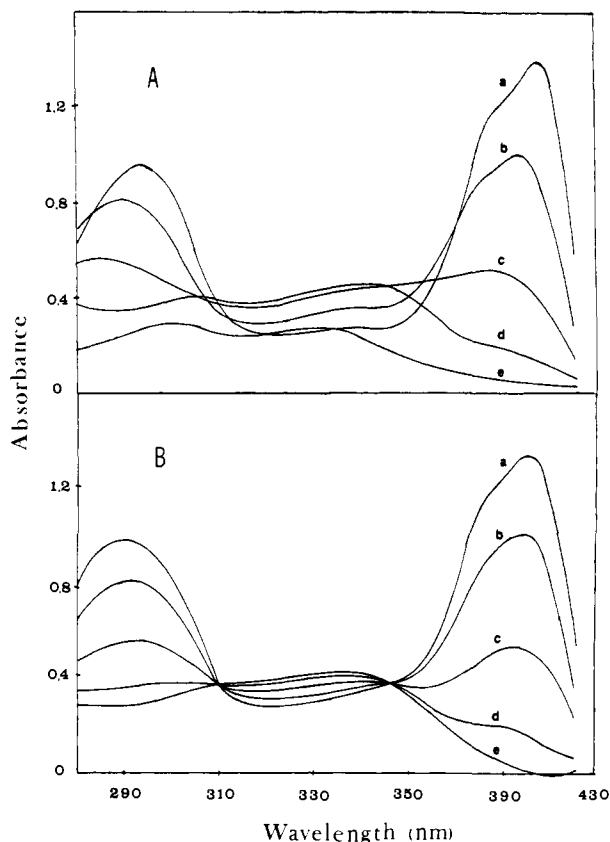
$$\log [XH^+]/[X] = -mH_0 + pK_{XH^+} \quad (4)$$

These are listed in Table II. Also listed in this table are  $m'$  and  $pK_{XH^+}$  values obtained by applying the similar equation:<sup>30</sup>

$$\log [XH^+]/[X] = -m'H_A + pK_{XH^+} \quad (5)$$

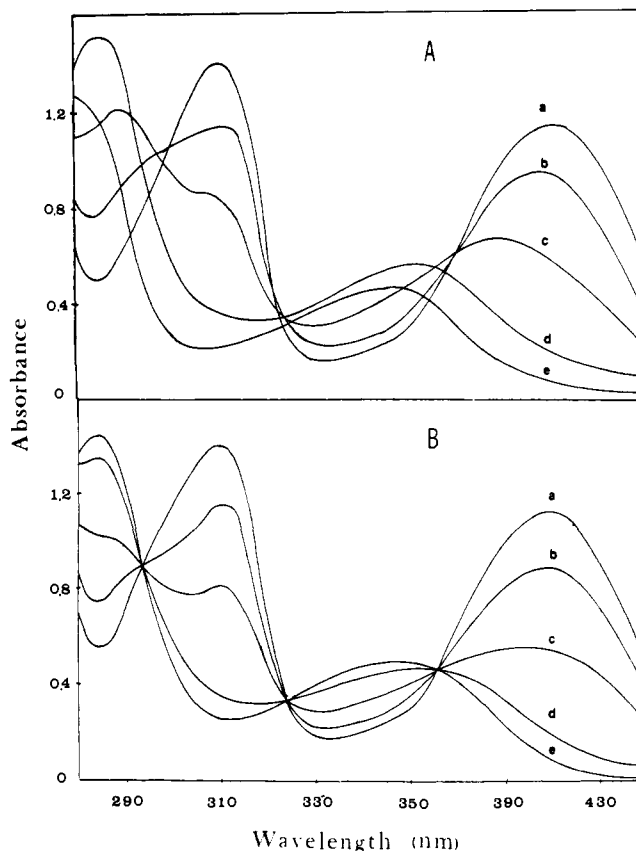
and  $\phi$  and  $pK_{XH^+}$  obtained by applying Bunnett and Olsen's equation:<sup>31</sup>

$$\log [XH^+]/[X] = -H_0 + \phi(H_0 + \log [H^+]) + pK_{XH^+} \quad (6)$$



**Figure 3.** Absorption curves of fluorenone in aqueous sulfuric acid: (a) 94.5%; (b) 83.5%; (c) 78.6%; (d) 73.0%; (e), 55.3%. Upper set: data curves. Lower set: curves reconstituted from the mean curve and first characteristic vector.

A necessary (but not sufficient) condition for affirming that a protonation reaction is governed by either  $H_0$  or  $H_A$  is that either  $m$  (eq 4) or  $m'$  (eq 5) is equal to unity. It can be seen that, as contended by Modena, Scorrano, and their coworkers,<sup>15</sup> no single acidity function suffices to describe the protonation of ketones. 2-Hydroxy-2-cyclohexen-1-one (the predominant tautomer of 1,3-cyclohexanedione<sup>32</sup>) and 2-cyclohexen-1-one follow  $H_A$ , as claimed by Zalewski and Dunn,<sup>8</sup> as does benzaldehyde; anthraquinone, one of Hammett and Deyrup's original indicators,<sup>33</sup> follows  $H_0$ ; acetophenone and benzophenone follow an acidity function intermediate between these two. In such cases we consider the  $pK_{XH^+}$  given

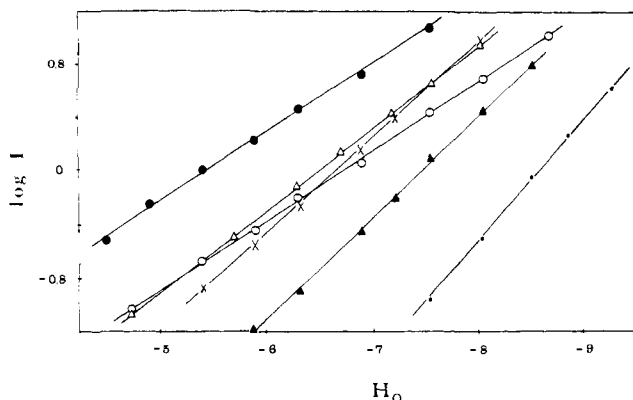


**Figure 4.** Absorption curves of anthraquinone in aqueous sulfuric acid: (a) 99.0%; (b) 91.6%; (c) 86.3%; (d) 80.7%, (e) 73.0%. Upper set: data curves. Lower set: curves reconstituted from the mean curve and first characteristic vector.

by the Bunnett-Olsen equation the most reliable value available at present. It can be observed that when either  $m$  or  $m'$  is close to unity, so that protonation is following either  $H_0$  or  $H_A$ , the  $pK_{XH^+}$  thus obtained is close to the Bunnett-Olsen  $pK_{XH^+}$  and deviates considerably from the  $pK_{XH^+}$  obtained using the alternative equation with  $m$  or  $m'$  deviating considerably from unity. For compounds such as acetophenone and benzophenone, for which both  $m$  and  $m'$  differ about equally from unity, and for which eq 4 should accordingly give too high a  $pK_{XH^+}$  and eq 5 too low a  $pK_{XH^+}$ , the Bunnett-Olsen equation yields an intermediate  $pK_{XH^+}$ , which we consider the best value for the moment.

**Table II.** Comparison of  $pK_{XH^+}$  Values Obtained by Equations 4, 5, and 6

Compd	Eq 4 ( $H_0$ )		Eq 5 ( $H_A$ )		Eq 6 (BO)	
	$m$	$-pK_{XH^+}$	$m'$	$-pK_{XH^+}$	$\phi$	$-pK_{XH^+}$
(1) 2-Hydroxy-2-cyclohexen-1-one	0.77	0.64	1.03	0.52	0.39	0.57
(2) 2-Cyclohexen-1-one	0.51	2.76	1.01	3.30	0.52	3.11
(3) Benzaldehyde	0.53	3.54	1.02	4.02	0.50	3.91
(4) Acetophenone	0.63	4.03	1.24	4.78	0.40	4.32
(5) Benzophenone	0.71	4.74	1.36	5.40	0.30	4.97
(6) Fluorenone	0.78	5.81	1.48	6.53	0.23	5.98
(7) Anthraquinone	0.94	8.08	1.44	7.30	0.06	8.13
(8) 2,4-Hexadienoic acid	0.46	2.88	0.90	3.32	0.57	3.30
(9) Benzoic acid	0.58	4.39	1.04	4.56	0.44	4.73
(10) Cinnamic acid	0.58	3.99	1.06	4.29	0.45	4.31
(11) Ethyl benzoate	0.74	5.94	1.24	5.85	0.27	6.16
(12) Ethyl cinnamate	0.72	4.83	1.38	5.56	0.30	5.06
(13) Acrylamide	0.34	1.00	0.50	0.86	0.83	1.15
(14) Benzamide	0.60	1.55	0.96	1.65	0.48	1.70
(15) <i>p</i> -Nitrobenzamide	0.59	1.95	0.98	2.13	0.48	2.13
(16) Cinnamamide	0.66	1.21	0.95	1.15	0.45	1.25



**Figure 5.** Plots of  $\log I$  ( $\equiv \log [XH^+]/[H^+]$ ) vs.  $H_0$  for benzaldehyde (O), 2-cyclohexen-1-one (●), acetophenone (Δ), benzophenone (X), fluorenone (▲), and anthraquinone (•).

The  $\phi$  and  $pK_{BH^+}$  values for 2-cyclohexen-1-one, benzophenone, benzamide (obtained by UV measurements), and acetophenone (obtained by NMR measurements) given by Scorrano et al.<sup>15</sup> are in fairly good agreement with those of Table II. However, many other values in the literature are in poor agreement. This is to be expected from their extreme variability.<sup>12,34</sup> Whenever  $pK_{XH^+}$  values obtained by UV spectroscopy deviate very widely from those of Table II, it is found that the spectra were subject to considerable solvent shifts and hence the  $pK_{XH^+}$  values must be regarded with suspicion.

The trends in  $m$  and  $\phi$  ( $\approx 1 - m$ ) shown in Table II can be interpreted qualitatively in terms of the "solvation hypothesis" of Scorrano and his co-workers.<sup>15</sup> When  $XH^+$  has the same solvation requirements as the protonated Hammett indicators  $BH^+$  used to establish the  $H_0$  scale,  $m = 1$  and  $\phi = 0$ ; when  $XH^+$  has the positive charge buried in a large and polarizable molecule, so that its solvation is less than that of  $BH^+$ ,  $m > 1$  and  $\phi$  is negative; when  $XH^+$  is a small and not very polarizable molecule, its solvation is greater than that of  $BH^+$ ,  $m < 1$ , and  $\phi$  is positive. In fact, all the compounds of Table II show positive  $\phi$  values, as expected for oxygen Bronsted bases<sup>15</sup> having conjugate acids with large solvation requirements.<sup>35</sup> However, among similar classes of compounds the smaller molecules show greater  $\phi$  values, or smaller  $m$  and  $m'$  values, indicative of greater solvation of  $XH^+$ .<sup>36</sup> This is particularly noticeable for the amides, acrylamide ( $\phi = 0.83$ ), benzamide ( $\phi = 0.48$ ), and cinnamide ( $\phi = 0.45$ ). The lower  $m$  (greater  $\phi$ ) values for Bronsted bases having a small molecular volume have been pointed out previously by Yates and McClelland<sup>35</sup> and explained by the smaller increase in  $f_X$  as the internal pressure increases with acid concentration; however, similar internal pressure changes should change  $f_{XH^+}$  equally, and the alternative explanation based on solvation effects would seem more plausible.

### Experimental Section

All of the carbonyl compounds were commercial products which were purified by distillation or recrystallization before use. Spectra

of sulfuric acid solutions at  $25.0 \pm 0.1$  °C were taken using procedures described previously.<sup>37</sup> Spectral curves were subjected to characteristic vector analysis, following a computer program described in ref 29. The ionization ratios and equilibrium constant were then calculated from the reconstituted spectra (cf. Figure 1B).

**Acknowledgment.** This work was supported by the National Research Council of Canada.

**Supplementary Material Available:** The computer program (ref 29) and ionization ratios of the bases in different concentrations of sulfuric acid (7 pages). Ordering information can be found on any current masthead page.

### References and Notes

- (1) L. A. Flexser, L. P. Hammett, and A. Dingwall, *J. Am. Chem. Soc.*, **57**, 2103 (1935).
- (2) L. P. Hammett, "Physical Organic Chemistry", McGraw-Hill, New York, N.Y., 1970, pp 290-297.
- (3) R. Stewart and K. Yates, *J. Am. Chem. Soc.*, **80**, 6355 (1958).
- (4) K. Yates and R. Stewart, *J. Am. Chem. Soc.*, **82**, 4059 (1960).
- (5) R. Stewart, M. R. Granger, R. B. Moodie, and J. J. Muenster, *Can. J. Chem.*, **41**, 1065 (1963).
- (6) R. Stewart and K. Yates, *Can. J. Chem.*, **37**, 664 (1959).
- (7) G. Culbertson and R. Pettit, *J. Am. Chem. Soc.*, **85**, 741 (1963).
- (8) R. I. Zalewski and G. E. Dunn, *Can. J. Chem.*, **46**, 2469 (1968).
- (9) E. M. Arnett, T. Cohen, A. A. Bothner-By., R. D. Bushick, and G. Sowinski, *Chem. Ind. (London)*, 473 (1961).
- (10) P. Haake, P. D. Cook, and G. H. Hurst, *J. Am. Chem. Soc.*, **89**, 2650 (1967).
- (11) K. Yates and H. Wai, *Can. J. Chem.*, **43**, 2131 (1965).
- (12) G. C. Greig and C. D. Johnson, *J. Am. Chem. Soc.*, **90**, 6453 (1968).
- (13) A. A. El-Anani, G. C. Greig, and C. D. Johnson, *Chem. Commun.*, 1024 (1970).
- (14) T. J. Gilbert and C. D. Johnson, *J. Am. Chem. Soc.*, **96**, 5846 (1974).
- (15) A. Levi, G. Modena, and G. Scorrano, *J. Am. Chem. Soc.*, **96**, 6505 (1974).
- (16) N. C. Deno, "Survey of Progress in Chemistry", Vol. 1, A. F. Scott, Ed., Interscience, New York, N.Y., 1963, p 155.
- (17) C. T. Davis and T. A. Geissman, *J. Am. Chem. Soc.*, **76**, 3507 (1954).
- (18) A. R. Katritzky, A. J. Waring, and K. Yates, *Tetrahedron*, **19**, 465 (1963).
- (19) C. D. Johnson, A. R. Katritzky, B. J. Ridgewell, H. Shakir, and A. M. White, *Tetrahedron*, **21**, 1055 (1965).
- (20) K. Yates and R. A. McClelland, *J. Am. Chem. Soc.*, **89**, 2686 (1967).
- (21) J. L. Simonds, *J. Opt. Soc. Am.*, **53**, 968 (1963).
- (22) R. L. Reeves, *J. Am. Chem. Soc.*, **88**, 2240 (1966).
- (23) T. W. Anderson, "Introduction to Multivariate Analysis", Wiley, New York, N.Y., 1958.
- (24) R. Harper, *Appl. Statistics*, **5**, 32 (1956).
- (25) Not all variables (e.g., the phases of the moon) are controlled, but many (apparently) are unimportant in the physical sciences.
- (26) P. H. Weiner, *J. Am. Chem. Soc.*, **95**, 5845 (1973), and references therein to earlier papers.
- (27) J. T. Bulmer and H. F. Shurwell, *J. Phys. Chem.*, **77**, 256 (1973).
- (28) S. Wold and M. Sjöström, *Chem. Scr.*, **2**, 49 (1972); S. Wold, *ibid.*, **5**, 97 (1974); M. Sjöström and S. Wold, *ibid.*, **6**, 114 (1974); S. Wold, *Acta Chem. Scand.*, **27**, 3602 (1973); S. Wold and K. Andersson, *J. Chromatogr.*, **80**, 43 (1973).
- (29) S. C. Wong, Ph.D. Thesis, McGill University, 1974.
- (30) In applying this equation we have corrected the  $H_A$  scale by increasing all values by 0.3; our unpublished researches have confirmed Bunnnett and Olsen's surmise<sup>31</sup> that the scale is incorrectly anchored in the dilute acid region.
- (31) J. F. Bunnnett and F. P. Olsen, *Can. J. Chem.*, **44**, 1899 (1966).
- (32) G. Schwarzenbach and E. Felder, *Helv. Chim. Acta*, **27**, 1044 (1944).
- (33) L. P. Hammett and A. J. Deyrup, *J. Am. Chem. Soc.*, **54**, 2721 (1932).
- (34) E. M. Arnett, R. P. Quirk, and J. W. Larsen, *J. Am. Chem. Soc.*, **92**, 3977 (1970).
- (35) K. Yates and R. A. McClelland, *Prog. Phys. Org. Chem.*, **2**, 323 (1974).
- (36) E. M. Arnett, "Proton Transfer Reactions", E. F. Caldin and V. Gold, Ed., Chapman and Hall, London, 1975, p 79.
- (37) W. I. Congdon and J. T. Edward, *J. Am. Chem. Soc.*, **94**, 6096 (1972).