

Determination of Positional Weighting Factors for the Swain and Lupton Substituent Constants \mathfrak{F} and \mathfrak{R}

S. G. Williams and F. E. Norrington*

Contribution from the Wellcome Research Laboratories, Beckenham, Kent, England. Received August 19, 1974

Abstract: Previous workers have shown that Hammett's substituent constant σ and related parameters can all be expressed as a linear combination of two principal components \mathfrak{F} and \mathfrak{R} , which represent respectively the field and resonance effects of the substituent without regard for its position in the molecule. A statistical analysis of data from 45 reaction series has now been carried out to determine positional weighting factors f and r for \mathfrak{F} and \mathfrak{R} , to enable the latter to be applied to the ortho, meta, and para positions of a benzene ring simultaneously. f_{para} and r_{para} were both defined as 1.000. f_{meta} was found to be 0.980 (not significantly different from 1.000), f_{ortho} 1.248, r_{meta} 0.347, and r_{ortho} 0.863. These values agree roughly with the ad hoc assumptions made hitherto by classical organic chemists and enable good correlations to be found between physico-chemical properties of the side chain and substituent parameters of mixed or multiply substituted populations of benzenoid compounds. They also demonstrate that generalized σ -type parameters may be successfully applied to the ortho position, although the correlation with observed data is less good than with the meta and para positions due, no doubt, to the proximity of the substituent to the side chain of the molecule.

Empirical relationships between structure and reactivity of organic molecules have fascinated physical chemists for many years. The best known equation is due to Hammett¹ and describes the influence of substituents in a benzene ring on reactions in the side chain. It may be formulated as

$$\log(k/k_0) = \rho\sigma \quad (1)$$

where k and k_0 are rate or equilibrium constants for the substituted and unsubstituted aromatic molecules, respectively, ρ is a reaction dependent proportionality constant, and σ is the substituent constant. For any substituent, these values are different at the meta and para positions. The σ constants, originally determined from the ionization of benzoic acids, were at first believed to be independent of the nature of the reaction but, as more experimental data were accumulated, they were found to be reaction dependent and not applicable to the general case.

To solve this problem, two approaches have been taken. The trivial approach has been to dismiss the concept of generally applicable, reaction independent substituent constants and to proliferate alternative sets of σ values each for a special system or application (for recent reviews, see Swain and Lupton²). A more exacting but far more useful solution has been attempted in extensions of the Hammett equation to a linear combination of fundamental components of the substituent constants.³

Classical thinking has recognized for many years that resonance effects due to conjugation of π electrons from the substituent with the aromatic π -electron cloud are mechanistically distinct from perturbations of σ electrons caused by the inductive effect of the electrostatic field associated with covalent bonds between atoms of differing electronegativity. It seemed plausible, therefore, to express substituent constants as a linear combination of those substituent parameters which were judged to measure, exclusively, the resonance or the inductive effects of a substituent. A substituent parameter (σ^*) has been derived from the inductive effects in aliphatic series.⁴ A suitably scaled version of this parameter, termed σ_I , also explains the inductive contribution to the observed polar effects of meta and para substituents on the ionization of benzoic acids^{5,6} and on the NMR lines of fluorobenzenes.⁷ Taft and Lewis³ therefore suggested the extended Hammett equation

$$\log(k^p/k_0) = \sigma_I + \sigma_R \quad (2)$$

to define the resonance effects (σ_R) of para substituents in the benzene ring, where σ_R is the difference ($\sigma - \sigma_I$), σ being derived via eq 1 from the ionization constants of para substituted benzoic acids. In its more general form, (3), where the subscript X refers to any substituent and σ , δ , and h are constants, this equation has been shown to correlate a wide variety of physical data.

$$\sigma_X = \lambda\sigma_{I,X} + \delta\sigma_{R,X} + h \quad (3)$$

In particular, it has been used by Charton⁸⁻¹¹ to demonstrate that the effects of ortho substituents are not dependent primarily upon steric effects but are essentially electrical in origin and also that such polar effects are not numerically equivalent for the ortho and para positions.

Swain and Lupton² made a fresh appraisal of the field and, invoking less assumptions than earlier studies, showed that a two-term equation similar to (3) explains, on the average, 93% of the variance in 42 different reaction series. In particular they demonstrated that a third parameter in (3) is superfluous, thus lending support to the classical concepts of the partitioning of polar effects.

The substituent parameters which emerged from the Swain and Lupton study were termed \mathfrak{F} and \mathfrak{R} and have the qualitative characteristics of σ_I and σ_R , respectively. They have the added virtues, however, of being independent of position in the aromatic ring, reaction, solvent, and temperature. Also they are available for 42 substituents,² and values for additions to the list may be simply derived from data on meta and para substituted benzoic acids (Table I).

When \mathfrak{F} and \mathfrak{R} were used instead of σ_I and σ_R with equation 3 and correlations obtained for the 42 sets of data, a wide range of λ/δ values was found.² This indicates that there is a broad continuum for the mixture of field and resonance effects which govern varied reaction series. On the other hand, Taft and Lewis³ have assumed that, although field and resonance parameters for a given substituent will vary if the position of the substituent is changed, the proportional change will be a constant for any two positions. They have quantitated this concept in the form of eq 4 and 5

$$\log(k^p/k_0) = \sigma_I + \sigma_R \quad (4)$$

$$\log(k^m/k_0) = \sigma_I + \frac{1}{3}\sigma_R \quad (5)$$

where the superscript p and m refer to para and meta substitution respectively. The ad hoc assumption that field ef-

fects are equal from the para and meta positions, whereas the resonance effect at the latter is reduced to $1/3$, has been shown to explain an overwhelming portion of the total variance in data from 43 varied reaction series.¹²

It seemed pertinent, therefore, to ask whether positional weighting factors for the parameters \mathcal{F} and \mathcal{R} could be derived by a systematic analysis of existing empirical data. Since the work of Charton⁸⁻¹¹ has shown that the ortho substituent effect is essentially polar in origin, unique weighting factors for the ortho position as well as meta and para have been sought. Such weighting factors should greatly widen the usefulness of \mathcal{F} and \mathcal{R} since physical properties of benzenoid systems substituted at different positions or even multiply substituted may then be correlated.

Analytical Methods. In the original Hammett equation for substituent effects (eq 1), if k and k_0 represent equilibrium or rate constants, then the left hand side of the equation is a free-energy related term. For any such term, P , then eq 1 may be generalized as

$$P = \rho\sigma + P^\circ \quad (6)$$

where P° refers to some standard or reference state. Swain and Lupton² have shown that, for any set of reaction data, i , that σ_i may be expressed as a linear combination of their substituent parameters \mathcal{F} and \mathcal{R} , i.e., for the k th substituent in the i th reaction set

$$P_{i,k} = \rho_i(\alpha'_i\mathcal{F}_k + \beta'_i\mathcal{R}_k) + P_i^\circ \quad (7)$$

where α'_i and β'_i are reaction dependent coefficients. This equation is formally equivalent to eq 2-5 which have been shown by many authors to correlate a wide variety of physicochemical data provided that all substituents k in a given reaction set i are substituted at the same position j in the parent, aromatic molecule. The hypothesis which the present study sets out to test is that position dependent terms f_j and r_j may be added to eq 7 which then enables data to be mixed and successfully correlated for reaction sets based on benzene derivatives substituted at either the ortho, meta, or para positions. To allow for uncertainties in the data, an error term e_i which is assumed to be normally distributed and independent of position and substituent is also added. Since ρ_i , α'_i , and β'_i are all purely set dependent, then the products $\rho_i\alpha'_i$ and $\rho_i\beta'_i$ in the expansion of (7) may be replaced by the set dependent coefficients α_i and β_i . The hypothesis to be tested can therefore be formalized as:

$$P_{ijk} = \alpha_i f_j \mathcal{F}_k + \beta_i r_j \mathcal{R}_k + P_i^\circ + e_{ijk} \quad (8)$$

For a given reaction series, the size of the coefficients α_i and β_i represents the sensitivity of the reaction to the substituent dependent products $f_j\mathcal{F}_k$ and $r_j\mathcal{R}_k$. If unique values of f_j and r_j can be determined, then α_i and β_i are readily found by multiple regression of P_{ijk} against $f_j\mathcal{F}_k$ and $r_j\mathcal{R}_k$.

In order to test eq 8, a matrix of data is required for i independent reaction series listing empirical, free-energy related parameters P_{ijk} for each of k substituents at each of j positions in the aromatic nucleus. Evaluation of the unknown terms in (8) then becomes a problem in statistical analysis. Because of the product terms $\alpha_i f_j$ and $\beta_i r_j$, the correlation is nonlinear and cannot be analyzed with standard multiple correlation procedures. The method which has been used for solving this problem is tolerant of gaps in the data matrix and is detailed in Appendix I. In essence, the analysis minimizes the sum of squares for each of the i reaction series fitted to eq 8 and evaluates the values of α_i , β_i , f_j , and r_j having adopted a normalizing convention that $f_{\text{para}} = r_{\text{para}} = 1.00$. The significance of the difference be-

Table I. Substituents and Substituent Constants

No.	Substituent	Formula	\mathcal{F}^a	\mathcal{R}^a
1	Acetylamino	NHCOCH ₃	0.470	-0.274
2	Acetoxy	OCOCH ₃	0.679	-0.071
3	Acetyl	COCH ₃	0.534	0.202
4	Acetylthio	SCOCH ₃	0.602	0.102
5	Amido	CONH ₂	0.437 ^b	0.079 ^b
6	Amino	NH ₂	0.037	-0.681
7	Benzamido	NHCOC ₆ H ₅	0.441 ^b	-0.167 ^b
8	Benzoyl	COC ₆ H ₅	0.481 ^b	0.189 ^b
9	Bromo	Br	0.727	-0.176
10	<i>n</i> -Butoxy	O(CH ₂) ₃ CH ₃	0.411	-0.551
11	<i>tert</i> -Butyl	C(CH ₃) ₃	-0.104	-0.138
12	Carboxy	COOH	0.552	0.140
13	Carboxylate anion	COO ⁻	-0.221	0.124
14	Chloro	Cl	0.690	-0.161
15	Cyano	CN	0.847	0.184
16	Diazonium cation	N ₂ ⁺	2.760	0.360
17	Dimethylamino	N(CH ₃) ₂	0.032 ^c	-0.863 ^c
18	Dimethyl S ⁺ cation	S(CH ₃) ₂ ⁺	1.678	-0.042
19	Ethoxy	OC ₂ H ₅	0.363	-0.444
20	Ethoxycarbonyl	CO ₂ C ₂ H ₅	0.552	0.140
21	Ethyl	C ₂ H ₅	-0.065	-0.114
22	Fluoro	F	0.708	-0.336
23	Formyl	CHO	0.667 ^b	-0.156 ^b
24	Hydrogen (unsubst)	H	0.000	0.000
25	Hydroxy	OH	0.487	-0.643
26	Hydroxymethyl	CH ₂ OH	0.129 ^b	0.011 ^b
27	Iodo	I	0.672	-0.197
28	Iodoxy	IO ₂	1.098	0.144
29	Mercapto	SH	0.464	-0.111
30	Methoxy	OCH ₃	0.413	-0.500
31	Methoxycarbonyl	COOCH ₃	0.552 ^d	0.140 ^d
32	Methyl	CH ₃	-0.052	-0.141
33	Methylseleno	SeCH ₃	0.221	-0.124
34	Methylsulfinyl	SOCH ₃	0.860	0.007
35	Methylsulfonyl	SO ₂ CH ₃	0.900	0.215
36	Methylthio	SCH ₃	0.332	-0.186
37	Nitro	NO ₂	1.109	0.155
38	<i>n</i> -Pentoxy	O(CH ₂) ₄ CH ₃	0.423	-0.577
39	Phenoxy	OC ₆ H ₅	0.747	-0.740
40	Phenyl	C ₆ H ₅	0.139	-0.088
41	Phosphonate anion	PO ₃ H ⁻	0.288	0.098
42	Isopropoxy	OCH(CH ₃) ₂	0.488	-0.724
43	<i>n</i> -Propoxy	O(CH ₂) ₂ CH ₃	0.369	-0.457
44	Isopropyl	CH(CH ₃) ₂	-0.080 ^e	-0.120 ^e
45	<i>n</i> -Propyl	(CH ₂) ₂ CH ₃	-0.070 ^e	-0.110 ^e
46	2-Siliconeopentyl	CH ₂ Si(CH ₃) ₃	-0.229	-0.081
47	Sulfamoyl	SO ₂ NH ₂	0.679	0.188
48	Sulfonate anion	SO ₃ ⁻	0.057	0.058
49	Trifluoromethyl	CF ₃	0.631	0.186
50	Trimethyl N ⁺ cation	N(CH ₃) ₃ ⁺	1.460	0.000
51	Trimethylsilyl	Si(CH ₃) ₃	-0.047	-0.044

^aUnless otherwise indicated, all the values are from ref 2. ^bCalculated from the σ_m and σ_p values listed in ref 17 using the equations $\sigma_m = 0.60\mathcal{F} + 0.27\mathcal{R}$ and $\sigma_p = 0.56\mathcal{F} + 1.00\mathcal{R}$ derived by Swain and Lupton.² ^cSame as for b but values taken from ref 42. ^dValue taken as for substituents 12 and 20. ^eValues interpolated between those for substituents 11, 21, and 32.

tween the values of f_j and r_j derived from all reaction sets and the values f_{ij} and r_{ij} derived from each of the individual reaction sets in turn was also determined statistically via Fisher's F ratio.

The Data Used. In order to minimize gaps in the data matrix referred to above it was essential to select a number of reaction series which included data for the ortho as well as for the meta and para positions. Attention was therefore given to the reviews of ortho-substituent effects given by Charton⁸ and Tribble and Traynham¹³ in addition to the reaction series listed by Swain and Lupton.² The original reaction data from these papers were tabulated and edited according to the following criteria. Firstly substituents were omitted for which values of \mathcal{F} and \mathcal{R} were not available (Table I). Secondly, substituents at the ortho position which had been shown by the above authors to deviate markedly

Table II. Reaction Series and Input Data

No.	Reaction ^a	P	No. of data points			Ref
			Ortho	Meta	Para	
Equilibria Data						
1	Dissociation of XC ₆ H ₄ COOH	σ	13	42	42	2, 17
2	Dissociation XC ₆ H ₄ COOH ^b	pK _a	6		7	32
3	Association XC ₆ H ₄ COOH with base ^c	Log K ^{'''}	11	12	11	19
4	Dissociation of XC ₆ H ₄ OH	σ ⁻	12	13	22	17
5	Dissociation XC ₆ H ₄ OH ^d	pK _a	9	6	14	31
6	Dissociation XC ₆ H ₄ (CH ₂) ₂ COOH	Log K _a	5	4	5	26
7	Dissociation XC ₆ H ₄ CH=CHCOOH	Log K _a	5	5	5	26
8	Dissociation XC ₆ H ₄ C≡CCOOH ^e	pK _a	6	6	6	28
9	Dissociation XC ₆ H ₄ C≡CCOOH ^f	pK _a	5	5	5	28
10	Basicity XC ₆ H ₄ NH ₂	pK _b	10	9	10	18
11	Basicity XC ₆ H ₄ NHNH ₂	pK _b	10	10	10	18
Kinetic Data						
12	Catalytic constant XC ₆ H ₄ COOH ^h	Log K _c	5	5	6	30
13	Esterification XC ₆ H ₄ C≡CCOOH ⁱ	Log k	5	5	5	28
14	Esterification XC ₆ H ₄ C≡CCOOH ^j	Log k	5	5	5	28
15	Esterification XC ₆ H ₄ C≡CCOOH ^k	Log k	4	4	4	27
16	Pyrolysis XC ₆ H ₄ CO ₂ CH(CH ₃) ₂ ^l	Log k	6	7	7	23, 24
17	Detritiation XC ₆ H ₄ C≡CD ^m	Log (k/k ^H) ^a	5	6	8	37
18	Hydrolysis XC ₆ H ₄ CH ₂ COOC ₂ H ₅ ⁿ	Log (k/k ^H) ^a		7	14	16
19	Solvolytic of XC ₆ H ₄ C(CH ₃) ₂ Cl ^o	σ ⁺		20	23	2
20	Hydrolysis XC ₆ H ₄ SO ₃ C ₆ H ₄ -p-CH ₃ ^p	Log k	7	4	4	20
21	Fission XC ₆ H ₄ CO ₂ (CH ₂) ₂ S(CH ₃) ₂ I ^q	Log k	5	5	5	25
22	Hydrolysis XC ₆ H ₄ OPO(OC ₂ H ₅) ₂ ^r	Log k	9	6	14	31
Infrared Spectral Data						
23	ν _{CO} for XC ₆ H ₄ CON(CH ₃) ₂ ^s	ν _{CO}		8	8	15
24	XC ₆ H ₄ CON(CH ₃) ₂ + C ₆ H ₅ OH ^t	Δν _{OH}		8	8	15
25	ν _{CO} for XC ₆ H ₄ CH=CHCON(CH ₃) ₂ ^s	ν _{CO}		8	8	15
26	XC ₆ H ₄ CH=CHCON(CH ₃) ₂ + C ₆ H ₅ OH ^t	Δν _{OH}		8	8	15
27	ν _{CO} for XC ₆ H ₄ CH=CHCOC ₆ H ₅ ^u	ν _{CO}	6	6	9	14
28	ν _{S=O} for p-CH ₃ C ₆ H ₄ SO ₃ C ₆ H ₄ X ^v	ν _{S=O}	7	4	3	22
29	ν _{OH} for XC ₆ H ₄ CH=NC ₆ H ₄ -o-OH ^v	Δν _{OH}	6	7	9	29
30	ν _{COP} for XC ₆ H ₄ OPO(OC ₂ H ₅) ₂ ^w	ν _{COP}	9	5	14	31
Nuclear Magnetic Resonance Data						
31	¹⁹ F NMR shift, XC ₆ H ₄ F ^x	σ ^F		15	15	2
32	¹ H NMR shift XC ₆ H ₄ OH ^y	Δδ	26	7	27	13
33	¹ H NMR shift XC ₆ H ₄ OCH ₃ ^z	δ	8	4	10	35
34	¹ H NMR shift XC ₆ H ₄ OCH ₂ CH ₃ ^{aa}	δ	4	4	5	35
35	¹ H NMR shift XC ₆ H ₄ OCH ₂ CH ₃ ^{bb}	δ	4	4	5	35
36	¹ H NMR shift XC ₆ H ₄ CH ₃ ^{cc}	δ	7	7	7	36
37	¹ H NMR shift XC ₆ H ₄ CH ₃ ^{dd}	δ	7	7	7	36
38	¹ H NMR shift XC ₆ H ₄ NH ₂ ^{ee}	δ	6	7	8	33
39	¹ H NMR shift XC ₆ H ₄ CHO ^{ff}	δ	5	7	10	34
Free Radical and Polarographic Data						
40	Free radical XC ₆ H ₄ CH ₂ ^{gg}	σ ^I		5	6	2
41	Oxidation of phenylferrocenes ^{hh}	ΔE _{1/4}	13	9	15	21
42	Reduction p-CH ₃ C ₆ H ₄ SO ₃ C ₆ H ₄ X ⁱⁱ	-E _{1/2}	7	4	4	38
43	Reduction p-CH ₃ C ₆ H ₄ SO ₃ C ₆ H ₄ X ^{jj}	-E _{1/2}	7	5	6	39
44	Reduction p-CH ₃ C ₆ H ₄ SO ₃ C ₆ H ₄ X ^{kk}	-E _{1/2}	7	5	6	39
45	Reduction p-CH ₃ C ₆ H ₄ SO ₃ C ₆ H ₄ X ^{ll}	-E _{1/2}	7	5	6	39

^a Based on acid dissociation constants (K_A) in water at 25° except where otherwise indicated. For series 1 and 4, P = equilibrium ($1/\rho$) log (K/K^H) values, for series 20, P = kinetic ($1/\rho$) log (K/K^H). Superscript H refers to the unsubstituted compound (X = H). For reaction 1, ρ = 1.00; 4, ρ = 2.23; 20, ρ = -4.54. ^b Dissociation constant (K_A) measured in 95% dimethyl sulfoxide in water at 25°. ^c Association constant (K'') determined spectrophotometrically in benzene at 25°. ^d Dissociation constant (K_A) determined in 0.1 M KCl at 25°. ^e Dissociation constant (K) measured in 35% dioxane in water at 25°. ^f As ^e, but in 50% aqueous ethanol. ^g Entry deleted in proofing. ^h The catalytic constant (k_c) for the dehydration of acetaldehyde hydrate in aqueous acetone at 25°, corrected for hydrogen ion catalysis. ⁱ Rate constant (k) for reaction with diphenyldiazomethane in absolute ethanol at 30°. ^j As ⁱ but in dioxane. ^k Rate constant (k) for esterification at 25° in methanol containing 0.4 M acid and 0.01 M HCl. ^l First-order rate constant (k) for the vapor phase elimination of propylene at 337.4°. ^m Rate of detritiation (k) in buffered 20% methanol-water, pH 8.05, 25°. ⁿ Rate constant (k) in alkaline, 60% acetone-water at 25°. ^o Solvolysis in 90% acetone-water at 25°. ^p Rate constant (k) determined in 60% ethanol-water at 60°. ^q Second-order rate constant (k) for hydrolysis in 0.1 N sodium hydroxide at 25°. ^r In CCl₄, 1 mg/ml, 34°. ^s As ^s but plus 0.2 mg/ml phenol. ^t In CHCl₃. ^u In CCl₄, 0.01 M, 20°. ^w Aryl-O-P stretching frequency (ν_{COP}) for neat compound. ^x Chemical shift (G , parts per million) in CCl₄ solution, $\sigma^F = G$. ^y As ^x in dimethyl sulfoxide. $\Delta\delta = \delta$ on scale where $\delta^H = 0$. ^z Chemical shift (G , parts per million) for methyl protons in CCl₄. ^{aa} Chemical shift (G , parts per million) for methylene protons, 10% in CCl₄. ^{bb} As ^{aa} but G for methyl protons. ^{cc} As ^z but in benzene. ^{dd} As ^z but in dioxane. ^{ee} Chemical shift (G) for amino protons in dimethyl sulfoxide. ^{ff} Solvent corrected chemical shift (G) for formyl proton in CCl₄ or CDCl₃. ^{gg} Benzyl free radicals ionized in gas phase by electron impact. Volts = σ^I . ^{hh} Chromapotentimetric quarter-wave potentials ($\Delta E_{1/4}$) relative to phenylferrocene measured in acetonitrile at 25°. ⁱⁱ Half-wave potential ($E_{1/2}$) for polarographic reduction in pyridine. ^{jj} As ⁱⁱ but in dimethylformamide. ^{kk} As ⁱⁱ but in dimethyl sulfoxide. ^{ll} As ⁱⁱ but in acetonitrile.

from their correlations were deleted. This ensured that, in general, the ortho data referred to the polar effects of the substituents as recognized by Charton and that gross anomalies due to H bonding or exceptionally bulky groups were omitted. The final compilation of 45 reaction series is listed in Table II together with a description of the reaction and source reference.^{2,13-39} In total 1050 data points are included, 279 at the ortho position and 335 and 436 at the meta and para positions, respectively.⁴⁰

Evaluation of f_j and r_j . Analysis of the data by the statistical methods given in Appendix I gave values for the positional weighting factors f_j and r_j which are listed below together with their standard errors:

$$\begin{aligned} \text{ortho position: } f_o &= 1.248 \pm 0.057, & r_o &= 0.863 \pm 0.048 \\ \text{meta position: } f_m &= 0.980 \pm 0.024, & r_m &= 0.347 \pm 0.028 \\ \text{para position: } f_p &= 1.000, & r_p &= 1.000 \end{aligned}$$

It is interesting to find that, although meta substituents are closer to the side chain reaction center than para substituents, the field effects from these positions are not significantly different. This has often been assumed a priori^{3,41} in previous attempts to factorize σ but was questioned by Swain and Lupton² who purposely avoided the need to make this assumption in their derivation of \mathcal{F} and \mathcal{R} . However, when the substituent is moved to the ortho position, closer still to the reaction center, the field effect is definitely increased. This result is qualitatively in accord with the increasing effect of substituents on the basicity or acidity of aliphatic amines or acids as the substituent is moved progressively nearer to the reaction center. In the aliphatic series, the substituent effect is approximately doubled for each step toward the α -carbon atom.⁴² Clearly in the aromatic system, the situation is quantitatively distinct.

The smallest value for any of the above resonance weighting factors is that for $r_m = 0.347$. This factor is closely analogous to the factor α used by Taft and Lewis³ in their attempt to factor σ into two principal components. They assumed that $f_m = f_p = 1.00$ and $r_m/r_p = \alpha$ and suggested that $\alpha = 0.333$. A statistical study of Roberts and Jaffé¹² using 43 reaction series concluded that an overwhelming portion of the total variance was accounted for with this value for α . The result from the present study of a group of 45 reactions agrees closely with this earlier work. None of the previous authors, however, proposed any value for the positional weighting factor at the ortho position. From this study, a value of $r_o = 0.863$ has been found which is near to, but significantly different from, the classical concept that $r_o = r_p = 1.000$ and consistent with the qualitative conclusions of Charton.

The Generality of f_j and r_j . The analysis made by Roberts and Jaffé also found, however, that, for some reaction series, a significant improvement in fit was obtained if the value of α was allowed to float. For comparison, the generality of the definitive weighting factors obtained in this study has been judged by using Fisher's F test to measure the significance of the differences $(f_j - f_{ij})$ and $(r_j - r_{ij})$ derived for all values of j at each value of i (see Appendix I). If the model tested is valid for each of the i reaction series, then the probability values derived from the F test should be rectangularly distributed between 0 and 1.0. The distribution obtained from the complete data is shown in Table III and, although essentially rectangular, there is a marked bias toward the low probability region. In fact, 7 of the 17 reaction series in the 0-10% probability band have probabilities less than 1%. These series are, therefore, likely to be outside the normal range of variation in the rest of the data. Nevertheless the remaining 38 reaction series (i.e., 84%) appear to fit the model.

A repeat of the analysis omitting all the ortho data gave

Table III. Frequency Distribution for the Complete Data Set

Probability %	Frequency ^a	% of total
0-10	17	37.8
10-20	5	11.1
20-30	2	4.4
30-40	2	4.4
40-50	2	4.4
50-60	5	11.1
60-70	2	4.4
70-80	4	8.9
80-90	2	4.4
90-100	4	8.9

^aTotal number reaction series is 45; ortho, meta, and para data included.

Table IV. Frequency Distribution for the Reduced Data Set

Probability %	Frequency ^a	% of total
0-10	5	11.6
10-20	3	7.0
20-30	2	4.6
30-40	5	11.6
40-50	3	7.0
50-60	1	2.3
60-70	11	25.6
70-80	4	9.3
80-90	2	4.6
90-100	7	16.3

^aTotal number reaction series is 43; meta and para data only.

the distribution in Table IV. It was also necessary to omit the benzoic acid series (No. 1, Table II) because \mathcal{F} and \mathcal{R} are derived directly from the meta and para data for this series² making the error terms zero. In addition series 2 was also omitted since meta data are not available in this case.

Using the remaining 43 series, the values of f_j and r_j were found to be:

$$\begin{aligned} \text{meta position: } f_m &= 1.002 \pm 0.017; & r_m &= 0.424 \pm 0.019 \\ \text{para position: } f_p &= 1.000 & r_p &= 1.000 \end{aligned}$$

The value of f_m is not significantly changed from that derived using all the data, but the value of r_m is slightly, but significantly, increased.

The distribution of probability values derived from Fisher's F test is shown in Table IV. Compared with Table III and considering the small numbers involved, this distribution is nicely rectangular. These results imply that the model defined in eq 8 is adequate for all the meta and para data but is deficient when certain reaction series with ortho substituents are included. Despite the conclusion of Charton,¹¹ it does seem, therefore, that neighboring group effects other than quantitative differences in field and resonance effects are significant in a minority of ortho cases.

Use of the Factors f_j and r_j . Swain and Lupton² stated explicitly that their parameters \mathcal{F} and \mathcal{R} were independent of reaction but, since they are derived from sets of data homogenous with respect to substituent position, then it is implicit that \mathcal{F} and \mathcal{R} are independent of this variable too. A simple test of the utility of the weighting factors derived above is, therefore, to see if their use improves the correlations between parameters \mathcal{F} and \mathcal{R} and measured physicochemical data for reaction series including a mixture of substituent positions. The results of the test applied to the 45 reaction series listed in Table II are given in Table V which compares the unweighted correlation

$$P_{ijk} = \alpha_i \mathcal{F}_i + \beta_i \mathcal{R}_k + P_i^\circ + e_{ijk} \quad (9)$$

with the weighted correlation, eq 8, first using all the data from Table II and then with the ortho data omitted.

The results of these two separate calculations are shown

Table V. Comparison of Weighted with Unweighted Correlations

Reaction no.	Ortho group included ^a				Ortho group omitted ^b			
	% explained variance		% improvement	No. of data points	% explained variance		% improvement	No. of data points
	Unweighted	Weighted			Unweighted	Weighted		
1	68.3	75.7	7.4	97	92.3	99.3	7.0	84
2	94.4	97.4	3.0	13	99.0	99.0	0.0	7
3	60.6	63.3	2.7	34	92.9	98.8	5.9	23
4	84.4	90.5	6.1	47	85.8	90.6	4.8	35
5	83.3	85.7	2.4	29	79.4	78.7	-0.7	20
6	83.2	83.7	0.5	14	96.8	98.8	2.0	9
7	87.5	91.6	4.1	15	85.6	94.4	8.8	10
8	82.2	81.7	-0.5	18	82.9	90.4	7.5	12
9	81.7	67.3	-14.4	15	97.8	97.3	-0.5	10
10	72.4	80.7	8.3	29	83.4	92.4	9.0	19
11	39.1	37.3	-1.8	30	37.7	36.6	-1.1	20
12	90.8	91.9	1.1	16	95.5	96.8	1.3	11
13	85.1	77.2	-7.9	15	88.8	87.0	-1.8	10
14	92.4	83.4	-9.0	15	95.9	97.0	1.1	10
15	76.4	69.4	-7.0	12	96.6	98.4	1.8	8
16	85.4	92.3	6.9	20	94.1	99.2	5.1	14
17	86.4	96.8	10.4	19	89.4	98.8	9.4	14
18 ^c	97.5	98.5	1.0	21	97.5	98.6	1.1	21
19 ^c	70.4	87.1	16.7	43	70.4	86.0	15.6	43
20	92.3	93.0	0.7	15	97.3	99.1	1.8	8
21	94.4	95.9	1.5	15	93.1	97.9	4.8	10
22	89.2	90.3	1.1	29	89.8	88.7	-1.1	20
23 ^c	87.0	95.6	8.6	16	87.0	95.6	8.6	16
24 ^c	91.7	97.4	5.7	16	91.7	97.6	5.9	16
25 ^c	81.6	84.2	2.6	16	81.6	85.0	3.4	16
26 ^c	76.3	88.2	11.9	16	76.3	87.4	11.1	16
27	87.5	95.7	8.2	21	86.5	95.5	9.0	15
28	95.5	94.1	-1.4	14	97.5	95.8	-1.7	7
29	91.6	94.4	2.8	22	91.5	97.6	6.1	16
30	31.5	35.7	4.2	28	43.1	40.8	-2.3	19
31 ^c	48.7	84.0	35.3	30	48.7	80.1	31.4	30
32	76.7	80.4	3.7	60	84.4	87.4	3.0	34
33	69.5	72.4	2.9	22	81.0	82.9	1.9	14
34	77.4	73.4	-4.0	13	97.7	94.1	-3.6	9
35	79.7	75.1	-4.6	13	96.4	94.2	-2.2	9
36	13.3	6.6	-6.7	21	87.8	86.7	-1.1	14
37	81.9	87.2	5.3	21	88.3	89.5	1.2	14
38	82.8	89.0	6.2	21	81.3	89.5	8.2	15
39	25.8	29.9	4.1	22	89.9	91.1	1.2	17
40 ^c	80.7	90.3	9.6	11	80.7	90.0	9.3	11
41	80.6	88.7	8.1	37	80.5	91.5	11.0	24
42	53.8	60.0	6.2	15	91.8	97.2	5.4	8
43	57.4	59.1	1.7	18	90.3	98.3	8.0	11
44	80.8	86.2	5.4	18	86.9	98.1	11.2	11
45	51.0	51.7	0.7	18	78.8	81.0	2.2	11
Mean:	75.6	78.9		Mean:	85.8	90.7		

^aFor weighted correlations: $f_o = 1.248$; $r_o = 0.863$; $f_m = 0.980$; $r_m = 0.347$; $f_p = r_p = 1.000$. ^bFor weighted correlations: $f_m = 1.002$; $r_m = 0.424$; $f_p = r_p = 1.000$. ^cReaction series including no data for ortho substituents.

in Table V and, as expected from the incompleteness of the model for ortho substituents, less variance is explained in their presence than in their absence. Nevertheless in both cases, the appropriate weighting factors improve the correlations. With ortho groups present, the percent explained variance is increased in 35 out of 45 cases, and the mean increases from 75.6 to 78.9%. Without ortho groups, the percent explained variance is again increased in 35 out of 45 cases and the mean increases from 85.8 to 90.7%. In general, therefore, use of the weighting factors will improve such correlations. The relatively good performance of the unweighted correlation results from the values for f_j and r_j all being close to unity except for r_{meta} . Nevertheless their use is necessary in principle and brings worthwhile benefits in practice.

Swain and Lupton² found for the 19 sets of experimental data used in their correlations vs. \mathcal{F} and \mathcal{R} that an average 93% of the variance was explained leaving only 7% attributable to experimental errors and failures of the theory. The reaction series studied by these authors were each consid-

ered to define a scale of substituent polar effects and were therefore measured with care, perhaps more so than for the present data which were taken more or less randomly from the literature. In the absence of ortho data, eq 8 leaves only 9% of the variance unexplained compared with the 7% of Swain and Lupton. Considering the degree by which their treatment has been extended, these additional errors are negligible. When ortho data are included, the unexplained variance rises to 21% which suggests that some 12% of the variance is then attributable to incompleteness of the model for series including ortho substituents.

Additivity of $f_j\mathcal{F}_k$ and $r_j\mathcal{R}_k$. The original polar substituent constants due to Hammett are additive over a polysubstituted aromatic system and can be used to predict accurately, for example, the pK_a values of polysubstituted benzoic acids or phenols¹⁷ and anilines⁴² (cf. eq 6). In an analogous manner, the pK_a values for polysubstituted members of these three series were calculated via eq 10 from \mathcal{F}_k and \mathcal{R}_k (Table I) and the values of f_j and r_j derived from all the data, i.e., inclusive of ortho substituents.

Table VI. Prediction of pK_a for Polysubstituted Benzoic Acids

No.	Substituents	Predicted pK_a^a	Measured pK_a^b	Difference
1	3,4-(CH ₃) ₂	4.36	4.41	-0.05
2	3,5-(CH ₃) ₂	4.29	4.30	-0.01
3	3,4,5-(OCH ₃) ₃	4.08	4.13	-0.05
4	3,4-(OH) ₂	4.25	4.49	-0.24
5	3,5-(OH) ₂	3.92	4.04	-0.12
6	3,4-(NO ₂) ₂	2.65	2.82	-0.17
7	3,5-(NO ₂) ₂	2.75	2.82	-0.07
8	4-CH ₃ -3,5-(NO ₂) ₂	2.89	2.97	-0.08
9	2,4-(CH ₃) ₂	4.43	4.18	0.25
10	2,5-(CH ₃) ₂	4.36	3.98	0.38
11	2,3-(OH) ₂	4.11	2.91	1.20
12	2,4-(OH) ₂	4.45	3.22	1.23
13	2,5-(OH) ₂	4.11	2.95	1.16
14	2-OH-4-NO ₂	3.43	2.23	1.20
15	2-OH-5-NO ₂	3.53	2.12	1.41
16	2,4-(NO ₂) ₂	2.41	1.42	0.99
17	2,5-(NO ₂) ₂	2.51	1.62	0.89
18	2-Cl-3-NO ₂	3.04	2.02	1.02
19	2-Cl-4-NO ₂	2.94	1.96	0.98
20	2-Cl-5-NO ₂	3.04	2.17	0.87
21	2,4-Cl ₂	3.45	2.68	0.77
22	2,5-Cl ₂	3.38	2.47	0.91
23	2,6-Cl ₂	3.33	1.59	1.74
24	2-OH-5-Cl	3.86	2.63	1.23
25	2-OH-5-Br	3.85	2.61	1.24
26	2,3-(CH ₃) ₂	4.36	3.74	0.62
27	2,6-(CH ₃) ₂	4.43	3.25	1.18
28	2,6-(CH ₃) ₂ -4-C(CH ₃) ₃	4.60	3.44	1.16
29	2,4,6-(CH ₃) ₃	4.57	3.44	1.13
30	2,3-(NO ₂) ₂	2.51	1.85	0.66
31	2,6-(NO ₂) ₂	2.27	1.14	1.13
32	2,4,6-(NO ₂) ₃	1.46	0.65	0.81
33	2-Cl-6-NO ₂	2.80	1.34	1.46
34	2-Br-6-NO ₂	2.78	1.37	1.41
35	2,6-(OH) ₂	4.31	1.08	3.23
36	2-OH-3-NO ₂	3.53	1.87	1.66
37	2-OH-6-NO ₂	3.29	2.24	1.05
38	2-OH-3,5-(NO ₂) ₂	2.82	0.70	2.12
39	2-OH-6-Cl	3.82	2.63	1.19
40	2-OH-6-CH ₃	4.37	3.32	1.05

^a Calculated from eq 11; see text. ^b Reference 17.

$$P_{\text{predicted}} = \alpha_i \sum f_j \mathfrak{F}_k + \alpha_i \sum r_j \mathfrak{R}_k + P_i^\circ + e_{ijk} \quad (10)$$

For each series, the values of α_i , β_i and P_i° were taken from the previously derived correlation (eq 8) between the measured parameter P_{ijk} and the products $f_j \mathfrak{F}_k$ and $r_j \mathfrak{R}_k$ using the data for the monosubstituted compounds which comprise the series listed in Table II. The relevant series are No. 1, 4, and 10 for benzoic acids, phenols, and anilines, respectively. In each case, the parameter P used for correlation was not pK_a but σ , σ^- , and pK_b , respectively, which, however, all bear exact linear relationships to pK_a as detailed in the review by Barlin and Perrin.¹⁷ Combination of the appropriate linear correlation (eq 8) with the expression for pK_a in terms of P_{ijk} gave the following predictive equations:

benzoic acids

$$pK_a(\text{predicted}) = 4.151 - 0.608 \sum_{j,k} f_j \mathfrak{F}_k - 0.823 \sum_{j,k} r_j \mathfrak{R}_k \quad (11)$$

phenols

$$pK_a(\text{predicted}) = 9.847 - 1.728 \sum_{j,k} f_j \mathfrak{F}_k - 2.777 \sum_{j,k} r_j \mathfrak{R}_k \quad (12)$$

anilines

$$pK_a(\text{predicted}) = 4.066 - 1.882 \sum_{j,k} f_j \mathfrak{F}_k - 4.148 \sum_{j,k} r_j \mathfrak{R}_k \quad (13)$$

Table VII. Prediction of pK_a for Polysubstituted Phenols

No.	Substituents	Predicted pK_a^a	Measured pK_a^b	Difference
1	3,4-(CH ₃) ₂	10.55	10.32	0.23
2	3,4-Cl ₂	8.09	8.59	-0.50
3	3,5-(CH ₃) ₂	10.30	10.15	0.15
4	3,5-Cl ₂	7.83	8.19	-0.36
5	3,5-(CH ₃) ₂ -4-NO ₂	7.94	8.24	0.30
6	3-C ₂ H ₅ -5-CH ₃	10.29	10.10	0.19
7	3,5-(CH ₃) ₂ -4-SO ₂ CH ₃	8.13	8.13	0.00
8	3,5-(CH ₃) ₂ -4-CN	8.33	8.21	0.12
9	3,4-(NO ₂) ₂	5.46	5.42	0.04
10	3,5-(OH) ₂	9.41	8.45	0.96
11	2,6-(CH ₃) ₂	10.76	10.59	0.17
12	2-CH ₂ OH-4-CH ₃	10.02	10.15	-0.13
13	2,6-(CH ₂ OH) ₂	9.24	9.66	-0.42
14	2,3-Cl ₂	7.74	7.70	0.04
15	2,5-Cl ₂	7.74	7.51	0.23
16	2,5-(NO ₂) ₂	5.08	5.22	-0.14
17	2,4,6-(NO ₂) ₃	2.00	0.42	1.58
18	2,4,6-(CH ₃) ₃ -3-NO ₂	9.21	8.98	0.23
19	2,4,5-(CH ₃) ₃	11.00	10.57	0.43
20	2,4,6-(CH ₃) ₃	11.23	10.88	0.35
21	2,4-(CH ₂ OH) ₂	9.12	9.77	-0.65
22	2,6-(CH ₂ OH) ₂ -4-CH ₃	9.71	9.92	-0.21
23	2,4-Cl ₂	8.00	7.85	0.15
24	2,6-Cl ₂	7.65	6.79	0.86
25	2,6-(NO ₂) ₂	4.36	3.71	0.65
26	2-OCH ₃ -5-CHO	9.14	8.89	0.25

^a Calculated from eq 12; see text. ^b Reference 17.

Table VIII. Prediction of pK_a for Polysubstituted Anilines

No.	Substituents	Predicted pK_a^a	Measured pK_a^b	Difference
1	3-OCH ₃ -5-NO ₂	1.74	2.13	-0.39
2	3,5-(OCH ₃) ₂	3.93	3.87	0.06
3	3-Br-4-CH ₃	3.65	4.06	-0.41
4	2,4,6-(NO ₂) ₃	-4.96	-9.40	4.44
5	2,4-(NO ₂) ₂	-1.82	-4.50	2.68
6	2,4-Cl ₂ -6-NO ₂	-0.74	-3.30	2.56
7	2-NO ₂ -4-Cl	0.30	-1.11	1.41
8	2,6-(CH ₃) ₂	5.32	3.97	1.35
9	2,3,5,6-(CH ₃) ₄	5.93	4.36	1.57
10	2,3-(CH ₃) ₂	5.00	4.77	0.23
11	2-OH-4,5-(CH ₃) ₂	6.22	5.26	0.96

^a Calculated from eq 13; see text. ^b Reference 42.

These equations were used to calculate the expected pK_a values for the multisubstituted compounds listed in Tables VI, VII, and VIII, and the results were compared with the published measured values.^{17,42}

Of the 40 substituted benzoic acids listed in Table VI, the first 8 do not contain ortho substituents and the differences between measured and predicted pK_a values are small. For the remaining 38 compounds, the differences are much larger and all positive, indicating that ortho groups consistently perturb the correlation. Regression of pK_a (predicted) vs. pK_a (measured) for all 40 acids gave a correlation in which only 57% of the variance is explained. Data listed in Table V suggest that, even with ortho substituents present, equations analogous to (8) should, in general, explain some 79% of the variance in such data. The failure of eq 11 implies, therefore, that either the additivity rule does not apply or that ortho substitution markedly changes the pK_a of benzoic acids by some additional mechanism.

As noted earlier, the size of the coefficients α_i and β_i represents the sensitivity of P_i to field and resonance effects. It was of interest, therefore, to compare the equation for the correlation of measured pK_a vs. $f_j \mathfrak{F}_k$ and $r_j \mathfrak{R}_k$ for benzoic acids all substituted in the ortho position, with that for benzoic acids without ortho substituents. Equation 10 was fitted first to the data for compounds 9-40 from Table

VI and then to the combined meta and para data from series 1, Table II giving eq 14 and 15 where the figures in parentheses are the t statistics for the coefficients.

ortho substituents present

$\text{p}K_a(\text{measured}) =$

$$-0.862 \sum_{j,k} f_j \mathcal{F}_k + 0.157 \sum_{j,k} r_j \mathcal{R}_k + 3.663$$

(10.52) (0.73) (14)

data points = 32, $F = 65.5$, explained variance = 82%

meta and para substituents only

$\text{p}K_a(\text{measured}) =$

$$-0.575 \sum_{j,k} f_j \mathcal{F}_k - 0.959 \sum_{j,k} r_j \mathcal{R}_k + 4.209$$

(74.32) (51.31) (15)

data points = 84, $F = 5492.7$, explained variance = 99%

The striking difference between these equations is that the coefficient of $r_j \mathcal{R}_k$ in (14) is very much smaller than in (15) and has a negligible t statistic, indicating that ortho substituents render the $\text{p}K_a$ of benzoic acids independent of resonance effects. Mechanistically this can be interpreted that ortho substituents perturb the coplanarity of the phenyl and carboxyl groups and inhibit the conjugation between their respective π -electron clouds and also explains the poor performance of (11) in predicting the $\text{p}K_a$ of ortho substituted benzoic acids. Nevertheless the high significance of (14) suggests that, for the polysubstituted acids 9-40 of Table VI, the field effects are essentially additive over all three substituent positions.

The regressions of $\text{p}K_a(\text{predicted})$ vs. $\text{p}K_a(\text{measured})$ for the heterogeneously substituted phenols and anilines listed in Tables VII and VIII explained 96 and 95% of the variance, respectively. These figures are as good as may be expected from the limitations of the theory and experimental measurements as discussed above. The additivity of \mathcal{F} and \mathcal{R} over all three substituent positions is therefore good for both of these series. In neither case is the conformation of the prosthetic group easily disturbed by ortho substituents.

Discussion

The model proposed to explain the polar effects of substituents in aromatic systems, which is embodied in eq 8, is mathematically straightforward and encompasses the classical ideas on the subject. Above all it is open to direct empirical test without involving any more than the very minimum of assumptions. A systematic analysis of this model by the statistical methods described above has shown that, for the great majority of reaction data from benzenoid compounds, unique and self consistent values exist for the substituent parameters \mathcal{F}_k and \mathcal{R}_k and the positional weighting factors f_j and r_j . Furthermore the products $f_j \mathcal{F}_k$ and $r_j \mathcal{R}_k$ are additive over a polysubstituted benzene ring so that, in the general case, adequate correlations can be achieved between these reaction independent constants and any free-energy related property of benzenoid compounds. Only in a minority of cases are such correlations inadequate and these anomalies appear to be related to neighboring group effects of ortho substituents. For benzoic acids, this effect was severe and probably due to steric inhibition of resonance of the carboxyl group whereas, for phenols and anilines, no such disturbance was noted. The essentially universal nature of these results makes possible, therefore, a description of substituent polar effects as simple and gener-

al as originally envisaged by Hammett. In particular it is a long way back from the "point of absurdity" noted by Swain and Lupton² when alternative sets of σ values for special systems and reactions were proliferating dramatically without increasing the convenience or validity of the approach.

The implications for such a widening of the scope of physicochemical vs. substituent parameter correlations are not only of interest to the physical chemist but are also considerable in the novel field of correlating biological effects of aromatic organic molecules with substituent effects. This field has been pioneered by Hansch⁴³ and, in a recent review,⁴⁴ the use of \mathcal{F} and \mathcal{R} in such correlations is discussed. Hansch considered it necessary to use eq 16

$$\log(k_X/k_H) = a_m \mathcal{F} + b_p \mathcal{F} + c_m \mathcal{R} + d_p \mathcal{R} + e \quad (16)$$

in order to mix data from meta and para substituted congeners. By analogy to define electronic effects in a series of congeners substituted at positions $j = 1, 2, 3$ (e.g., ortho, meta, and para in a benzenoid compound) eq 17 would be needed.

$$P_{ijk} = \alpha_{i,j=1} \mathcal{F}_k + \alpha_{i,j=2} \mathcal{F}_k + \alpha_{i,j=3} \mathcal{F}_k + \beta_{i,j=1} \mathcal{R}_k + \beta_{i,j=2} \mathcal{R}_k + \beta_{i,j=3} \mathcal{R}_k + e_{ijk} \quad (17)$$

Each of the coefficients α_{ij} and β_{ij} could be determined empirically by a multiple regression vs. p_{ijk} . This approach requires a considerable number of data points for P_{ijk} and was ruled out by Hansch on practical grounds since the necessary number of compounds in a congeneric series would rarely be available. The present method, however, allows eq 10 to be used in place of (17) with some confidence. Only coefficients α_i and β_i need to be estimated which can be achieved with no more than 8 or 10 suitable congeners.

There is a fundamental distinction between the statistical solutions to the problem initiated by Swain and Lupton,² and carried further in this paper, and the rigorous physicochemical description of substituent effects. A multicomponent analysis of the latter can yield the greatest amount of information⁴⁵ but is of less use to the practicing organic or biological chemist than a generalized statistical approach on the lines of the original Hammett equation.

It is of interest that the method adopted by Swain and Lupton did not limit the description of substituent polar effects to two components. They looked for and expected to find that at least three parameters would be necessary to explain their empirical data but, in fact, found that two were both necessary and sufficient. Since these parameters were found to be qualitatively in accord with the σ_I and σ_R scales of Taft and Lewis,³ it is natural to call them field (\mathcal{F}) and resonance (\mathcal{R}) effects, although no precise meaning attaches to these terms. Caution should be exercised, therefore, in interpreting the results of such correlations in detailed mechanistic terms. However, the parameter \mathcal{F} is defined on the mechanistically reasonable assumption that $\mathcal{F} \equiv \sigma$ for the 4-substituted [2.2.2]bicyclooctyl-1-carboxylic acids and \mathcal{R} is defined similarly on the assumption that $\mathcal{R} \equiv 0$ for the trimethylammonium cation. It is of interest, therefore, that recent calculations of charge densities⁴⁶ by Pople's CNDO/2 method have shown that the substituent effects on the fluorine $2p_y(\sigma)$ orbital charge density of para substituted fluorobenzenes and of 4-substituted [2.2.2]bicyclooctyl 1-fluorides are equal and correspond well to the σ_I scale. Theoretical support is therefore provided for the idea that the statistical partitioning of substituent polar effects is consistent with that expected from classical chemical concepts.

Finally one may speculate that the same systematic method as has been used to quantify the weighting factors f_j

and r_j for the benzene ring may be applied to polycyclic or heteroaromatic structures with equal effect. It is known that Hammett σ values can predict the pK_a 's of naphthoic acids, pyridines, etc.,^{17,42} so no doubt the model represented by eq 8 will apply to such systems.

Acknowledgments. We are grateful to Messrs. D. Field and G. C. Sheppey for performing the mathematical analysis. We also thank Mr. O. Fanimo and Miss J. Matthews for technical assistance and Drs. P. J. Goodford, R. M. Hyde, and R. Wootton for helpful discussions.

Appendix I (by D. Field and G. C. Sheppey)

Mathematical Model. Let P_{ijk} be the measurement for the compound of series i with substituent k in position j . Then the model which is assumed is

$$P_{ijk} = P_i^0 + \alpha_i f_j \mathfrak{F}_k + \beta_i r_j \mathfrak{R}_k + e_{ijk} \quad (\text{A1})$$

where \mathfrak{F}_k and \mathfrak{R}_k are the known values of \mathfrak{F} and \mathfrak{R} for substituent k .

f_j and r_j are the factors defining the positional effect of \mathfrak{F} and \mathfrak{R} for position j and are defined to be unity for the para position. α_i , β_i , and P_i^0 are constants for series i . e_{ijk} is the measurement error and is assumed to be normally distributed with mean zero and variance σ_i^2 for series i .

Assuming this model and given data of P_{ijk} for various i , j , and k , the constants f_j and r_j , and also α_i , β_i , and P_i^0 , together with the variances σ_i , can be estimated by maximum likelihood.

The likelihood (L) is defined as

$$L = \prod_{ijk} \frac{1}{\sigma_i (2\pi)^{1/2}} \times \exp[-(P_{ijk} - P_i^0 - \alpha_i f_j \mathfrak{F}_k - \beta_i r_j \mathfrak{R}_k)^2 / 2\sigma_i^2] \quad (\text{A2})$$

whence

$$\log L = -\sum_{ijk} \log \sigma_i + \frac{1}{2} \log (2\pi) + \frac{(P_{ijk} - P_i^0 - \alpha_i f_j \mathfrak{F}_k - \beta_i r_j \mathfrak{R}_k)^2}{2\sigma_i^2} \quad (\text{A3})$$

The maximum of L is equivalent to the maximum of $\log L$, and conditions for a maximum can be obtained by equating derivatives to zero. Thus

$$\frac{\partial \log L}{\partial \alpha_i} = \frac{1}{2\sigma_i^2} \sum_{jk} f_j \mathfrak{F}_k \times (P_{ijk} - P_i^0 - \alpha_i f_j \mathfrak{F}_k - \beta_i r_j \mathfrak{R}_k) = 0 \quad (\text{A4})$$

$$\frac{\partial \log L}{\partial \beta_i} = \frac{1}{2\sigma_i^2} \sum_{jk} r_j \mathfrak{R}_k \times (P_{ijk} - P_i^0 - \alpha_i f_j \mathfrak{F}_k - \beta_i r_j \mathfrak{R}_k) = 0 \quad (\text{A5})$$

$$\frac{\partial \log L}{\partial P_i^0} = \frac{1}{2\sigma_i^2} \sum_{jk} (P_{ijk} - P_i^0 - \alpha_i f_j \mathfrak{F}_k - \beta_i r_j \mathfrak{R}_k) = 0 \quad (\text{A6})$$

$$\frac{\partial \log L}{\partial \sigma_i} = \sum_{jk} -1/\sigma_i + (P_{ijk} - P_i^0 - \alpha_i f_j \mathfrak{F}_k - \beta_i r_j \mathfrak{R}_k)^2 / \sigma_i^3 = 0 \quad (\text{A7})$$

Equations A4–A6 are equivalent to linear regression for each series separately with P_{ijk} as dependent variable and $f_j \mathfrak{F}_k$ and $r_j \mathfrak{R}_k$ as independent variables.

Equation A7 can be rearranged to

$$n_i \sigma_i^2 = S_i = \sum_{jk} (P_{ijk} - P_i^0 - \alpha_i f_j \mathfrak{F}_k - \beta_i r_j \mathfrak{R}_k)^2 \quad (\text{A8})$$

where n_i is the number of measurements in series i , and S_i is the sum of squares of residuals from the regression of series i referred to above.

Substitution from eq A8 into eq A3 gives

$$\log L = -N \log (2\pi) / 2 - \sum_i (n_i \log \sigma_i + n_i / 2)$$

i.e.,

$$\log L = -N(\log (2\pi) + 1) / 2 - \sum_i n_i \log \sigma_i$$

i.e.,

$$\log L = -N(\log (2\pi) + 1) / 2 - \sum_i n_i \log (S_i / n_i) / 2$$

i.e.,

$$\log L = -N(\log (2\pi) + 1) / 2 + \left(\sum_i n_i \log n_i \right) / 2 - \left(\sum_i n_i \log S_i \right) / 2 \quad (\text{A9})$$

where $N = \sum_i n_i$ is the total number of measurements.

Thus in order to find a maximum of $\log L$, it is sufficient to find a minimum of

$$M = \sum_i n_i \log S_i \quad (\text{A10})$$

since the other terms in eq A9 are constant and M appears with a negative sign.

Given a set of values of f_j and r_j for all positions j , it is possible to calculate the regressions for each series, thus calculate all S_i , and therefore calculate M from eq A10. The values of f_j and r_j which give a minimum of M can be found by a suitable nonlinear minimization procedure such as that due to Powell⁴⁷ which has been used in this paper.

Testing the Model. A more general model than eq A1 is

$$P_{ijk} = P_i^0 + a_{ij} \mathfrak{F}_k + b_{ij} \mathfrak{R}_k + e_{ijk} \quad (\text{A11})$$

and eq A1 is a particular case of this in which the hypothesis

$$a_{ij} = \alpha_i f_j \text{ and } b_{ij} = \beta_i r_j \quad (\text{A12})$$

is assumed.

The validity of this hypothesis can be tested by analysis of variance for each series i provided that the values of f_j and r_j which are assumed are estimated without reference to the data for series i . However, if there are data for a large number of series, the analysis of variance will be approximately true even though the data for series i are used to estimate f_j and r_j . Thus

$$\phi_i = \left(\left(S_i - \sum_j t_{ij} \right) / (3m_i - 3) \right) / \left(\sum_j t_{ij} / (n_i - 3m_i) \right)$$

will be approximately distributed like Fisher's F ratio with $(3m_i - 3)$ and $(n_i - 3m_i)$ degrees of freedom, where t_{ij} is the sum of squares of residuals for the regression of series i and position j with P_{ijk} as dependent variable and \mathfrak{F}_k and \mathfrak{R}_k as independent variables, and m_i is the number of positions for which there are data for series i .

Accordingly if p_i is the probability that the value ϕ_i will be exceeded in Fisher's F distribution with $(3m_i - 3)$ and $(n_i - 3m_i)$ degrees of freedom, then the set of values of p_i for all series should be uniformly distributed between 0 and 1.

Supplementary Material Available: a complete tabulation of the reaction data (21 pages). Ordering information is given on any current masthead page.

References and Notes

- (1) L. P. Hammett, *Chem. Rev.* **17**, 125 (1935).
- (2) C. G. Swain and E. C. Lupton, *J. Am. Chem. Soc.*, **90**, 4328 (1968).
- (3) R. W. Taft, Jr., and I. W. Lewis, *J. Am. Chem. Soc.*, **80**, 2436 (1958).

- (4) R. W. Taft, *J. Am. Chem. Soc.*, **75**, 4231 (1953).
 (5) J. D. Roberts and W. T. Moreland, Jr., *J. Am. Chem. Soc.*, **75**, 2167 (1953).
 (6) M. S. Newman, "Steric Effects in Organic Chemistry", Wiley, New York, N.Y., 1956, Chapter 13, pp 578-580, 594-597.
 (7) R. W. Taft, Jr., *J. Am. Chem. Soc.*, **79**, 1045 (1957).
 (8) M. Charton, *J. Am. Chem. Soc.*, **91**, 6649 (1969).
 (9) M. Charton and B. I. Charton, *J. Org. Chem.*, **36**, 260 (1971).
 (10) M. Charton, *J. Org. Chem.*, **36**, 266 (1971).
 (11) M. Charton, *J. Org. Chem.*, **36**, 882 (1971).
 (12) J. L. Roberts and H. H. Jaffe, *J. Am. Chem. Soc.*, **81**, 1635 (1959).
 (13) M. Tribble and J. G. Traynham, *J. Am. Chem. Soc.*, **91**, 379 (1969).
 (14) N. Silver and D. W. Boykin, Jr., *J. Org. Chem.*, **35**, 759 (1970).
 (15) K. Spaargaren, C. Kruk, T. A. Molensar-Langeveld, P. K. Korver, P. J. van der Haak, and Th. J. de Boer, *Spectrochim. Acta, Part A*, **28**, 965 (1972).
 (16) Y. Yukawa, Y. Tsuno, and M. Sawada, *Bull. Chem. Soc. Jpn.*, **39**, 2274 (1966).
 (17) G. Barlin and D. D. Perrin, *Q. Rev., Chem. Soc.*, **20**, 75 (1966).
 (18) H. H. Stroh and G. Westphal, *Ber.*, **96**, 184 (1963).
 (19) M. M. Davis and H. B. Hetzer, *J. Res. Nat. Bur. Stand.*, **60**, 569 (1958).
 (20) V. M. Maramae and J. R. Asenbush, *Org. React.*, **2**, 83 (1965).
 (21) W. F. Little, C. N. Reilley, J. D. Johnson, K. N. Lynn, and A. P. Sanders, *J. Am. Chem. Soc.*, **86**, 1376 (1964).
 (22) V. M. Maramae, *Org. React.*, **4**, 34 (1967).
 (23) G. G. Smith and D. A. K. Jones, *J. Org. Chem.*, **28**, 3496 (1963).
 (24) D. A. D. Jones and G. G. Smith, *J. Org. Chem.*, **29**, 3531 (1964).
 (25) P. Marmalis and H. N. Rydon, *J. Chem. Soc.*, 1049 (1955).
 (26) J. F. J. Dippy and J. E. Page, *J. Chem. Soc.*, 357 (1938).
 (27) M. S. Newman and S. H. Merrill, *J. Am. Chem. Soc.*, **77**, 5552 (1955).
 (28) I. J. Solomon and R. Filler, *J. Am. Chem. Soc.*, **85**, 3492 (1963).
 (29) A. W. Baker and A. T. Shulgin, *J. Am. Chem. Soc.*, **81**, 1523 (1959).
 (30) R. P. Bell and W. C. E. Higginson, *Proc. R. Soc. London, Ser. A*, **197**, 141 (1949).
 (31) C. van Hooendonk and L. Ginjaar, *Recl. Trav. Chim. Pays-Bas*, **86**, 449 (1967).
 (32) M. Hojo, M. Utaka, and Z. Yoshida, *Kogyo Kagaku Zasshi*, **69**, 885 (1966).
 (33) B. M. Lynch, B. C. Macdonald, and J. G. K. Webb, *Tetrahedron*, **24**, 3595 (1968).
 (34) R. E. Klinck and J. B. Stothers, *Can. J. Chem.*, **40**, 1071 (1962).
 (35) C. Heathcock, *Can. J. Chem.*, **40**, 1865 (1962).
 (36) N. Nakagawa and S. Fujiwara, *Bull. Chem. Soc. Jpn.*, **34**, 143 (1961).
 (37) C. Eaborn, G. A. Skinner, and D. R. M. Walton, *J. Chem. Soc. B*, 922 (1966).
 (38) V. M. Maramae, *Org. React.*, **5**, 953 (1968).
 (39) V. M. Maramae, *Org. React.*, **4**, 573 (1967).
 (40) For the complete tabulation of these data, see paragraph at the end of this paper concerning supplementary material.
 (41) R. W. Taft, Jr., and I. C. Lewis, *J. Am. Chem. Soc.*, **81**, 5343 (1959).
 (42) J. Clark and D. D. Perrin, *Q. Rev., Chem. Soc.*, **18**, 295 (1964).
 (43) C. Hansch and T. Fujita, *J. Am. Chem. Soc.*, **86**, 1616 (1964).
 (44) C. J. Cavallito, "Structure-Activity Relationships", Vol. I, Pergamon Press, New York, N.Y., 1973, Chapter 3, p 101.
 (45) S. Ehrenson, R. T. C. Brownlee, and R. W. Taft, *Prog. Phys. Org. Chem.*, **10**, 1 (1973).
 (46) R. T. C. Brownlee and R. W. Taft, *J. Am. Chem. Soc.*, **92**, 7007 (1970).
 (47) M. J. D. Powell, *Comput. J.*, **7**, 155 (1964).

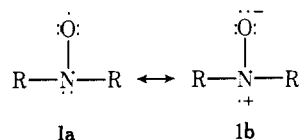
The Ease of Formation of Thionitroxide Radicals¹

Wayne C. Danen* and David D. Newkirk

Contribution from the Department of Chemistry, Kansas State University, Manhattan, Kansas 66506. Received February 20, 1975

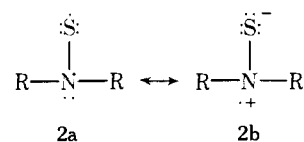
Abstract: The thermal decomposition of bis(pyrrolidyl-1) disulfide (**8**) was studied over the temperature range 45–68° by a radical scavenging technique utilizing Banfield's radical. The activation parameters obtained ($\Delta H^\ddagger = 30.8$ kcal/mol and $\Delta S^\ddagger = 6.5$ eu) indicated a relatively weak S–S bond in bis(dialkylamino) disulfides, $R_2NS-SNR_2$, and appreciable stability for thionitroxide radicals, $R_2NS\cdot$. The thermal and photolytic decompositions of several bis(dialkylamino) disulfides and bis(diphenylamino) disulfide were investigated by electron spin resonance spectroscopy and spectra of the corresponding thionitroxides observed. Phenyl radicals were found to react approximately twice as fast with bis(pyrrolidyl-1) disulfide as with cyclopentyl disulfide. The small difference in reactivity was interpreted in terms of a stepwise, addition-elimination displacement process with a rate-limiting formation of a metastable sulfur intermediate and little or no S–S bond homolysis in the transition state.

Most types of free radicals are transient species which are reactive with other compounds or are prone to self-reaction via dimerization or disproportionation. If, however, these self-reactions are prohibited by steric effects, or if the radical is stabilized by electronic effects, a free radical sufficiently stable to isolate and store may result. Several classes of such radicals are known,² with the nitroxides³, **1**,



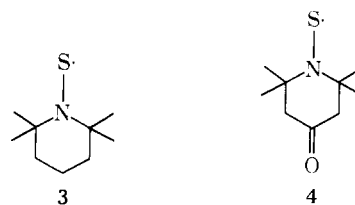
being one of the most common types of neutral radicals. Nitroxides have an inherently stable electronic arrangement about the oxygen and nitrogen atoms, and most organic nitroxides exhibit little tendency to dimerize at this center at room temperature.⁴

The remarkable stability of the nitroxide group has prompted us to investigate the sulfur analogue of these radicals, i.e., disubstituted aminothiyl radicals or thionitroxides, **2**. A priori, thionitroxides are expected to be less stable than the corresponding oxygen analogues because of the decreased importance of the charge-separated resonance **2b** relative to **1b** as a result of the lower electronegativity of



sulfur and the relatively long N–S bond. These expectations are borne out since, unlike nitroxides, thionitroxides exist as dimers at room temperature, as attested to by the fact that many bis(disubstituted amino) disulfides, R_2NSSNR_2 , have been synthesized and isolated.

The existence of the thionitroxides **3** and **4**, generated



from the corresponding disulfides by heating at 90–200°C, has been demonstrated by electron spin resonance (ESR) spectroscopy by Bennett, Sieper, and Tavs.⁸ Both **3** and **4** gave simple three-line ESR spectra: $a^N = 11.4$ G and 10.9 G; $g = 2.0173$ and 2.0171, respectively. These parameters