Phosphonate Complexes. Part 5.† An Improved Multiparametric Refinement Program, MUCOMP, for Complex Formation Studies from Potentiometric Data

By **Michel Wozniak,*** **Joël Canonne,** and **Guy Nowogrocki,** Ecole Nationale Supérieure de Chimie de Lille University of Lille 1, B.P. 108, 59652 Villeneuve d'Ascq, France

A new version of the refinement program MUCOMP is described. The stability constants of the complex species as well as the parameters liable to comport systematic errors (total concentration of main components and of impurities, and electrodes characteristics) are adjusted so as to minimize the weighted error-square sum of the added volume, the partial derivatives with respect to the unknown parameters being evaluated analytically. A generalized Newton–Raphson method is employed for the simultaneous calculation of the free concentrations and of the added volume. To deal with carbonate-contaminated reagents and the subsequent carbon dioxide evolution which occurs during titrations under nitrogen flow, an appropriate function for carbon dioxide departure is introduced in the proton balance and the corresponding equilibrium defect is taken into account in the weighting procedure.

pH TITRATION techniques are convenient for investigation of complex formation between metal ions and ligands which are conjugate bases of weak acids. We have previously reported a multiparametric refinement program, MUCOMP, for the computation of stability constants of complexes $M_qH_jA_p$ (M = metal ion, A = ligand) from potentiometric data corresponding to base titrations of complexant–cation mixtures.[1] To avoid any systematic error, these stability constants are refined together with the other significant parameters, which are (i) the analytical concentrations of the components (ligand, metal ion, strong acid or strong base, carbonate) and (ii) the electrode characteristics (zero shift and slope).

The most troublesome of the impurities is the carbonic acid introduced by contaminated reagents. In conventional titrations, performed with a nitrogen flow over or through the solution, a variable amount of the resulting carbon dioxide is carried away and thus prevents strict application of the mass-balance equations: the removal of carbon dioxide, which is efficient from acidic solutions, becomes negligible from neutral or basic medium. This behaviour can partly explain the existence of 'hysteresis' between direct and reverse titrations. The first purpose of this paper is to apply a function for the carbon dioxide evolution in the proton balance which has been successfully used in simple acid–base neutralizations.[2]

In addition, the iterative Newton–Raphson method has been generalized to calculate simultaneously the free concentrations of the components A and M, and the added volume. In fact, the habitual procedure where the added volume is calculated apart from the free concentrations is deficient when this volume does not appear only within the dilution term, *e.g.* when ligand or metal ion is added. This rigorous approach can readily be modified to deal with other titration modes and leads to obtention of the analytical expressions of the derivatives of the added volume with respect to the unknown parameters.

This work arises from studies of complex formation between metal ions and a number of phosphonates. Therefore, only the theoretical background of the program will be developed in the present paper; full calculation details and complete results will be published in the following descriptive parts.

RESULTS AND DISCUSSION

Mass-balance Equations.—The titration starts with a volume v_0 of a solution containing one ligand (total concentration C_A^0, number of introduced protons N_A), one metal ion (C_M^0), a strong protolyte ($H_i^0 > 0$ for a strong acid, $H_i^0 < 0$ for a strong base) and absorbed carbonate (C_i^0). The addition of a volume v_n of strong base ($H_x^0 < 0$, total concentration of the alkali-metal ion) containing carbonate as impurity (C_x^0) should theoretically lead to the mass-balance equations (1)—(3) where

$$\frac{v_0}{v_0 + v_n} C_A^0 = a + \Sigma p \beta_{qjp} m^q h^j a^p \qquad (1)$$

$$\frac{v_0}{v_0 + v_n} C_M^0 = m + \Sigma q \beta_{qjp} m^q h^j a^p \qquad (2)$$

$$\frac{v_0}{v_0 + v_n} (H_i^0 + N_A C_A^0 + 2C_i^0) + \frac{v_n}{v_0 + v_n} (H_x^0 + 2C_x^0) = h + \Sigma j \beta_{qjp} m^q h^j a^p + 2[H_2CO_3] + [HCO_3^-] \qquad (3)$$

a, m, and h are the free concentrations and $\beta_{qjp} = [M_qH_jA_p](m^qh^ja^p)^{-1}$ are stoicheiometric overall stability constants. Evidently, a positive H_x^0 value corresponds to a reverse titration with a strong acid.

The third equation is the hydrogen balance where the species H_2CO_3 and HCO_3^-, coming from the contaminated solutions, appear. However, in titrations performed under an inert gas flow, the total amount of carbonate remaining in solution is smaller than expected (below),

$$(v_0 + v_n)([H_2CO_3] + [HCO_3^-] + [CO_3^{2-}]) < v_0C_i^0 + v_nC_x^0$$

i.e. some CO_2 being carried away. In a previous work,[2] a semi-empirical function for the CO_2 evolution has been proposed and tested on typical acid–base neutralizations. For reproducible experimental conditions (increment of volume, maximum waiting time, stirring, gas purging *etc.*) we have supposed that the proportion of CO_2 leaving

---

† Part 4, M. Wozniak and G. Nowogrocki, *Talanta*, 1979, **26**, 1135.

the solution, G_n, is essentially dependent on the H_2CO_3 concentration in solution (see below; β_1 and β_2 are the over-

$$G_n = \frac{\alpha\beta_2 h^2}{\varepsilon\beta_2 h^2 + \beta_1 h + 1}$$

all stability constants of H_2CO_3, α and ε being constants at given experimental conditions). In acidic medium G_n tends toward α/ε, but for the sake of simplification, we have considered that the CO_2 departure must then be quantitative ($\alpha = \varepsilon$). At an experimental point n, that function is applied to the remaining carbonate in solution, to which has been added an increment $(v_n - v_{n-1})C_x{}^0$. Thus, the carbonate balance becomes that shown below which yields

$$(v_0 + v_n)([H_2CO_3] + [HCO_3{}^-] + [CO_3{}^{2-}]) = [v_0\boldsymbol{M}_n C_i{}^0 + (v_n - \boldsymbol{F}_n)C_x{}^0](1 - G_n)$$

$$\text{with} \quad \begin{cases} \boldsymbol{M}_n = \Pi(1 - G_n) \text{ and } M_1 = 1 \\ \boldsymbol{F}_n = \boldsymbol{F}_{n-1} + G_{n-1}(v_{n-1} - \boldsymbol{F}_{n-1}) \text{ and } \boldsymbol{F}_1 = 0 \end{cases}$$

$$\frac{v_0}{v_0 + v_n}(H_i{}^0 + N_A C_A{}^0) + \frac{v_n}{v_0 + v_n} H_x{}^0 + \frac{1}{v_0 + v_n}[v_0\boldsymbol{M}_n C_i{}^0 + (v_n - \boldsymbol{F}_n)C_x{}^0](1 - G_n)(2 - \bar{n}_H) = h + \Sigma j\beta_{qjp}m^q h^j a^p \quad (4)$$

a more realistic proton balance, equation (4), where $\bar{n}_H$ is the average number of protons bound to carbonate ion, the free OH^- concentration being included in the summation ($\beta_{0-10} = K_w$). Equation (4), where a dynamic system takes place, is valuable only for reproducible experimental conditions.

The hydrogen ion concentration h, appearing in the previous equations, is calculated from the e.m.f. of the cell by equation (5). Here, E represents the true

$$-\log h = E_0 + Pj_a h + PK_w j_b h^{-1} + PE \quad (5)$$

potential or the pH, E_0 is a constant for the electrodes used, j_a and j_b are liquid junction coefficients, P is a slope including the sensitivity of the glass electrode ($P \approx 1$ for pH readings and $P \approx 1/0.059$ for potential readings).

$$\begin{vmatrix} \boldsymbol{C}_A - (a + \Sigma p\boldsymbol{c}_k) \\ \boldsymbol{C}_M - (m + \Sigma q\boldsymbol{c}_k) \\ \boldsymbol{C}_H - (h + \Sigma j\boldsymbol{c}_k) \end{vmatrix} = \begin{vmatrix} \Sigma p^2\boldsymbol{c}_k & \Sigma pq\boldsymbol{c}_k & \frac{v_{nc}}{v_0 + v_{nc}}\boldsymbol{C}_A \\ \Sigma pq\boldsymbol{c}_k & \Sigma q^2\boldsymbol{c}_k & \frac{v_{nc}}{v_0 + v_{nc}}\boldsymbol{C}_M \\ \Sigma pj\boldsymbol{c}_k & \Sigma qj\boldsymbol{c}_k & \frac{v_{nc}}{v_0 + v_{nc}}[\boldsymbol{C}_H - H_x{}^0 - C_x{}^0(1 - G_n)(2 - \bar{n}_H)] \end{vmatrix} \times \begin{vmatrix} \frac{\Delta a}{a} \\ \frac{\Delta m}{m} \\ \frac{\Delta v_{nc}}{v_{nc}} \end{vmatrix} \quad (6)$$

Calculation of a, m, and v_{nc}.—The first step in the treatment of the data is the calculation, for each experimental point n, of the free concentrations a,m and of the added volume v_{nc}, from given values of β_{qjp} and h. In the earlier version, as in a number of refinement programs (GAUSS,[3] SCOGS,[4] LETAGROP,[5] DALSFEK [6]), the values of a and m were calculated by a Newton–Raphson iteration in the mass-balance equations (1) and (2), where the dilution term was obtained from the experimental added volume v_n. It was then possible to calculate the volume of the added reagent v_{nc} from the proton balance (3). However, as pointed out in the introduction, this treatment is wrong in principle, for the added volume appearing in equations (1) and (2) must be considered unknown, since the residuals are calculated from v_n or from its transformed function C_H. The MINIQUAD [7,8] approach is basically different since a, m, and β_{qjp} are corrected simultaneously, but for given values of h and v_n.

We have then extended the iterative Newton–Raphson method to relation (3) in order to obtain a, m, and v_{nc} simultaneously. A first-order Taylor expansion about the starting set of guessed values gives a system of linear equations which can be written in the matrix notation (6) where $\boldsymbol{C}_A$, $\boldsymbol{C}_M$, $\boldsymbol{C}_H$ are the left members of expressions (1), (2), and (4) with $v_n = v_{nc}$, and $\boldsymbol{c}_k = \beta_{qjp}m^q h^j a^p$. The elements of the matrix are the analytical expressions of the derivatives with respect to ln a, ln m, and ln v_{nc}.

The solution of this system by matrix inversion yields the relative shifts $\Delta a/a$, $\Delta m/m$, $\Delta v_{nc}/v_{nc}$ and leads to the new values $a \exp(\Delta a/a)$, $m \exp(\Delta m/m)$, and $v_{nc}(1 + \Delta v_{nc}/v_{nc})$; exponentiation is used for the free concentrations to overcome negative values in the beginning of the process. To avoid misbehaviour on not too good guesses, these relative corrections are maximized to +2.3 and −2.3. The iterations are repeated until the fractional shifts become smaller than 10^{-7}. The starting values of a and m for each point, except for the first one, are the values found for the preceding point. For the first experimental point, a preliminary calculation of a and m is performed using only equations (1) and (2), from a first guess corresponding to $a = C_A{}^0$ and $m = C_M{}^0$.

At first sight, that extension seems to be superfluous. Effectively, when the error-square sum (see next paragraph) is small, $v_n \approx v_{nc}$, and the results obtained from the classical two-steps procedure are identical to those of the generalized approach. One of the main advantages of the latter method is in its applicability to various titration procedures by means of appropriate corrections. For example, when the ligand is added to a solution of metal ion, $v_0C_A{}^0$ is just substituted by $v_nC_{Ax}{}^0$ ($C_{Ax}{}^0$ = total concentration of the added ligand) in equations (1) and (4). In that case, the multiparametric refinement would be applied to the analysis of the solution rather than to the determination of stability constants. This method could be used for more complicated systems by introducing the appropriate mass-balance equations.

Refinement of the Parameters.—The conventional Gauss–Newton least-squares method is used to evaluate the shifts to be applied at the refined parameters. The weighted error-square sum to be minimized is given by equation (7) where N is the total number of experimental

$$S = \sum_{n=1}^{N} W_n(v_n - v_{nc})^2 \qquad (7)$$

points and W_n a weighting factor which includes the possible sources of experimental errors,[2] equation (8), where

$$\frac{1}{W_n} = \sigma_{0v}{}^2 + \sigma_{0E}{}^2\left(\frac{\partial v_{nc}}{\partial E}\right)^2 + k^2\sigma_{0E}{}^2\left(\frac{\beta_1 h}{\beta_2 h^2 + \beta_1 h + 500}\right)^2 \qquad (8)$$

$\sigma_{0v}{}^2$ and $\sigma_{0E}{}^2$ are the variances on v and E, the last term being an estimation of the error due to the potential drift caused by the CO_2 evolution.[2] The values of σ_{0v}, σ_{0E}, and k are chosen so as to obtain, in the absence of any systematic error, a variance of residuals ($s^2 = S$/degrees of freedom) in the range of unity. By differentiating S with respect to each parameter $\boldsymbol{p}_r$ and equating to zero, we obtain a set of R linear equations (R being the number of refined parameters) in respect to the relative corrections $\Delta\boldsymbol{p}_r/\boldsymbol{p}_r$, equation (9).

$$\sum_{n=1}^{N} W_n(v_n - v_{nc})\frac{\partial v_{nc}}{\partial \ln \boldsymbol{p}_r} = \sum_{n=1}^{N}\sum_{u=1}^{R} W_n \frac{\partial v_{nc}}{\partial \ln \boldsymbol{p}_r}\frac{\partial v_{nc}}{\partial \ln \boldsymbol{p}_u}\frac{\Delta \boldsymbol{p}_u}{\boldsymbol{p}_u} \qquad (9)$$

Relative shifts have been introduced by transforming the derivatives into $\frac{\partial v_{nc}}{\partial \boldsymbol{p}_r}\,.\,\boldsymbol{p}_r = \frac{\partial v_{nc}}{\partial \ln \boldsymbol{p}_r}$, to handle terms of the same order of magnitude. As in the previous version,[1] the partial derivatives are evaluated analytically. For parameters appearing in equations (1), (2), and (4), their general expression is given by equation (10) where

$$\frac{\partial v_{nc}}{\partial \ln \boldsymbol{p}_r} = v_{nc}\cdot\boldsymbol{p}_r\left[\frac{\partial(\boldsymbol{C}_A - \Sigma p\boldsymbol{c}_k)}{\partial \boldsymbol{p}_r}\,\mathbf{s}[3,1] + \frac{\partial(\boldsymbol{C}_M - \Sigma q\boldsymbol{c}_k)}{\partial \boldsymbol{p}_r}\,\mathbf{s}[3,2] + \frac{\partial(\boldsymbol{C}_H - \Sigma j\boldsymbol{c}_k)}{\partial \boldsymbol{p}_r}\,\mathbf{s}[3,3]\right] \qquad (10)$$

$\mathbf{s}[i,j]$ are elements of the inverse of the matrix (6). The Table gives the particular expressions of the necessary derivatives, which must be taken only with respect to a parameter $\boldsymbol{p}_r$. Furthermore, in the Table the coefficients corresponding to $\frac{\partial v_{nc}}{\partial \ln h}$ have been included.

The derivatives with respect to E_0, P, and E, the latter being introduced in W_n, are obtained from equation (5) by applying the relation $\frac{\partial v_{nc}}{\partial \ln X} = X\cdot\frac{\partial v_{nc}}{\partial \ln h}\cdot\frac{\partial \ln h}{\partial X}$. However, in consideration of the little influence of Pj_a and PK_wj_b and of the practical constancy of P and K_w, the following variable change $J_a = Pj_a$ and $J_b = PK_wj_b$ was made in equation (5) to simplify the calculations. It must be pointed out that J_a and J_b take different values depending on whether potential ($P \approx 1/0.059$) or pH ($P \approx 1$) is introduced. This leads to the expressions shown below.

$$\frac{\partial v_{nc}}{\partial \ln E_0} = -\frac{E_0}{1/\ln 10 + J_a h - J_b h^{-1}}\frac{\partial v_{nc}}{\partial \ln h}$$

$$\frac{\partial v_{nc}}{\partial \ln P} = -\frac{PE}{1/\ln 10 + J_a h - J_b h^{-1}}\frac{\partial v_{nc}}{\partial \ln h}$$

$$\frac{\partial v_{nc}}{\partial E} = -\frac{P}{1/\ln 10 + J_a h - J_b h^{-1}}\frac{\partial v_{nc}}{\partial \ln h}$$

The processed data correspond in fact to several titrations performed under different experimental conditions, *i.e.* with variable total concentrations and electrode responses. So as in the LETAGROP approach,[5,9] the adjusted parameters fall in two classes: (i) the stability constants ($\boldsymbol{p}_r = \beta_{qjp}$) which are common to all the data and (ii), the parameters particular to each titration curve ($\boldsymbol{p}_r = C_A{}^0$, $C_M{}^0$, $H_i{}^0$, $H_x{}^0$, $C_i{}^0$, $C_x{}^0$, α, E_0, P) which have significance only for a certain part of the data used. Then, in the set of equations (9), $\partial v_{nc}/\partial \ln \boldsymbol{p}_r$ will be non-vanishing only (i) if $\boldsymbol{p}_r = \beta_{qjp}$ or (ii) if the experimental

Coefficients appearing in the analytical expressions of the derivatives [equation (10)]

$\boldsymbol{p}_r$	$\frac{\partial(\boldsymbol{C}_A - \Sigma p\boldsymbol{c}_k)}{\partial \boldsymbol{p}_r}$	$\frac{\partial(\boldsymbol{C}_M - \Sigma q\boldsymbol{c}_k)}{\partial \boldsymbol{p}_r}$	$\frac{\partial(\boldsymbol{C}_H - \Sigma j\boldsymbol{c}_k)}{\partial \boldsymbol{p}_r}$
$C_A{}^0$	$\frac{v_0}{v_0 + v_{nc}}$	0	$\frac{v_0}{v_0 + v_{nc}} N_A$
$C_M{}^0$	0	$\frac{v_0}{v_0 + v_{nc}}$	0
$H_i{}^0$	0	0	$\frac{v_0}{v_0 + v_{nc}}$
$H_x{}^0$	0	0	$\frac{v_{nc}}{v_0 + v_{nc}}$
$C_i{}^0$	0	0	$\frac{v_0\boldsymbol{M}_n}{v_0 + v_{nc}}(1 - G_n)(2 - \bar{n}_H)$
$C_x{}^0$	0	0	$\frac{v_{nc} - \boldsymbol{F}_n}{v_0 + v_{nc}}(1 - G_n)(2 - \bar{n}_H)$
α	0	0	$\frac{2 - \bar{n}_H}{v_0 + v_{nc}}\left\{\left(v_0C_i{}^0\frac{\partial \boldsymbol{M}_n}{\partial \alpha} - C_x{}^0\frac{\partial \boldsymbol{F}_n}{\partial \alpha}\right)(1 - G_n) - [v_0\boldsymbol{M}_nC_i{}^0 + (v_{nc} - \boldsymbol{F}_n)C_x{}^0]\frac{\partial G_n}{\partial \alpha}\right\}$
β_{qjp}	$-p\,.\,m^qh^ja^p$	$-q\,.\,m^qh^ja^p$	$-j\,.\,m^qh^ja^p$
h	$-\Sigma jp\boldsymbol{c}_kh^{-1}$	$-\Sigma jq\boldsymbol{c}_kh^{-1}$	$-\Sigma j^2\boldsymbol{c}_kh^{-1} - 1 - \frac{v_0\boldsymbol{M}_nC_i{}^0 + (v_{nc} - \boldsymbol{F}_n)C_x{}^0}{v_0 + v_{nc}}\left[(2 - \bar{n}_H)\frac{\partial G_n}{\partial h} + (1 - G_n)\frac{\partial \bar{n}_H}{\partial h}\right]$

point n belongs to the particular titration for which the parameter $\boldsymbol{p}_r$ is applicable. The matrix design takes a particular form which has been previously described.[1]

The correction vector is obtained by the conventional matrix resolution and gives, in most cases, a set of improved values which are used as the estimates for the next calculation. The corresponding square sum is lowered and after a finite number of iterations a minimum value of S is reached. The final values are then printed.

This strategy, where common and particular parameters are adjusted at the same time, is very powerful in compensating systematic errors. Therefore the selection of the species which is founded on the agreement between experimental and calculated data is improved: unrealistic complexes are sometimes introduced only to compensate neglected systematic errors. In our refinements, the 'goodness of fit' was estimated from the variance of the residuals which must be close to unity.

The Program MUCOMP.—A new version of the MUCOMP * program based on the present work has been written in HPL language, akin to BASIC, for the Hewlett-Packard 9825 calculator (24 kbytes memory). The use of this laboratory calculator makes it possible to automate fully the collection of titration data,[2,10] *i.e.* added volume of reagent and potential, which are stored on magnetic tape. The processing of these data can then be performed immediately after the titrations.

Besides the values of all the parameters, the usual output of the program [1] furnishes the variance of the weighted residuals, the standard deviations (σ) on the refined parameters, and the matrix of the correlation coefficients. These latter values can be used for the estimation of the errors associated with the stepwise constants obtained from cumulative constants.[11,12] Furthermore, this standard output has been improved by plotting, on the HP 9862 A plotter, some typical curves [titration curve, distribution of the residuals, $\sqrt{W_n} = f(v_n)$].

A detailed account of the calculations will be given in further papers devoted to the determination of stability constants of phosphonate complexes. Particular attention will be paid to (i) the weighting procedure, (ii) the choice of the parameters to be refined, and (iii) the model testing.

In such studies the MUCOMP approach proves to be very useful, since (i) the CO_2 evolution often interferes, (ii) some of the investigated compounds being rather unstable, the solutions are made just before titration by weighing out small quantities of ligand and the refinement of its total concentration becomes essential, and (iii) the compensation of systematic errors makes it possible to check minor species.

[1/609 *Received*, 14*th April*, 1981]

---

* A listing of the MUCOMP program will be sent on request.

REFERENCES

[1] M. Wozniak and G. Nowogrocki, *Talanta*, 1978, **25**, 643.

[2] M. Wozniak and G. Nowogrocki, *Talanta*, 1981, **28**, 575.

[3] R. S. Tobias and M. Yasuda, *Inorg. Chem.*, 1963, **2**, 1307.

[4] I. G. Sayce, *Talanta*, 1968, **15**, 1397.

[5] R. Arnek, L. G. Sillen, and O. Wahberg, *Ark. Kemi*, 1969, **31**, 353.

[6] R. M. Alcock, F. R. Hartley, and D. E. Rogers, *J. Chem. Soc., Dalton Trans.*, 1978, 115.

[7] A. Sabatini and A. Vacca, *J. Chem. Soc., Dalton Trans.*, 1972, 1693.

[8] A. Sabatini, A. Vacca, and P. Gans, *Talanta*, 1974, **21**, 53.

[9] L. G. Sillen and B. Warnqvist, *Ark. Kemi*, 1969, **31**, 315.

[10] G. Nowogrocki, J. Canonne, and M. Wozniak, *Anal. Chim. Acta*, 1979, **112**, 185.

[11] M. Wozniak, D. Thesis, University of Lille (France), 1977.

[12] M. Micheloni, A. Sabatini, and A. Vacca, *Inorg. Chim. Acta*, 1977, **25**, 41.