General Digital Computer Program for Chemical Equilibrium Calculations

Chemical equilibrium problems, especially those involving ionic equilibria, occupy a large portion of the undergraduate chemistry curriculum. The modern approach to the solution of equilibrium problems usually involves writing the pertinent equilibrium constant expressions along with additional equations specifying the conditions of the particular problem, such as material balance equations, the electroneutrality equation or the proton balance equations, and solving these for the unknown concentrations. Although writing equations specifying the problem is straightforward, the actual solution of these non-linear simultaneous equations often presents difficulties. At the elementary level, these equations are usually solved by making approximations which will make the equations either linear or particularly well-behaved non-linear ones. Unfortunately, as problems become more complex making a priori reasonable approximations becomes more difficult. Other approaches to the solution of these equations involve graphical (1) or numerical methods. The numerical methods have received repeated attention [see (2) and the references contained therein], and generally involve algebraic manipulation of the equations to obtain one or more higher degree equations in a few unknowns, which are then solved using well-known numerical methods such as the Newton-Raphson or iterative procedures (3).

The aim of the present work was to show how a general digital computer program for solving equilibrium problems could be formulated. Use of a systematic trial-and-error numerical method allowed the equilibrium constant and material and charge balance equations to be written in the program in a form closely corresponding to the equations as initially written in the statement of the problem. A program in this form does not require tedious algebraic combinations of equations; this allows a general program in which only the equations to be solved, the values of the constants and analytical concentrations, and the output format must be specified for any given problem. The method used is described below. The program, written in a commonly used computer language, FORTRAN, is given and explained through the use of two examples.

Formulation of the Method

Since a computer is incapable of operating in terms of unknown variables, algebraic equations as such cannot be introduced into the computer. The approach taken then to solve $n$ algebraic equations in $n$ unknowns is to guess values for two variables, which are called, by analogy to graphical methods (1), the master variables. The remaining variables are then calculated using $n - 2$ of the equations written in only slightly modified form. This modified form for writing the $n - 2$ equations consists of solving each for one of the variables in terms of the others, and sequencing the equations in such a manner that any variable can be computed using the master variables and variables already computed in the sequence. The remaining two equations are used as tests. If both are satisfied when all of the variables are substituted into them, the problem is solved. If they are not satisfied, the master variables are systematically varied until the two test equations are satisfied.

The master variables [called $CX$ and $CY(I)$ in the program] are varied in the following manner. $CX$ is initially given the guessed value and two $CY$ values are found, $CY(1)$ and $CY(2)$, each of which satisfy one of the two test equations. If the correct $CX$ has not been used, $CY(1)$ will not equal $CY(2)$. $CX$ is then varied by a systematic procedure, new values of the variables and $CY(1)$ and $CY(2)$ are calculated, and this procedure is continued until $CY(1)$ equals $CY(2)$. A flow chart of the method is given in Figure 1. Although the

![Figure 1. Overall flow chart of computer program for solving equilibrium problems.](image-url)
systematic trial-and-error procedure used in this program is somewhat slower than iterative or the Newton-Raphson methods, it is more generally applicable and does not suffer the convergence problems associated with these methods (5). The time required on a high speed digital computer is often negligibly different. The actual procedure of carrying out these operations, the method of making initial guesses for the master variables, and the manner in which the program operates is best illustrated by the solution of a simple problem.

Example 1.

Calculate the concentrations of all species in a CF solution a weak acid, H₂A.

1) Formulation of Equations. The six species in this system are H⁺, OH⁻, H₂A, H₃A⁻, HA²⁻, and A³⁻. They are related through the following equations:

\[ K_1 = \frac{[H₂A⁻][H⁺]}{[H₂A]} \]
\[ K_2 = \frac{[HA²⁻][H⁺]}{[H₃A⁻]} \]
\[ K_3 = \frac{[A³⁻][H⁺]}{[HA²⁻]} \]
\[ K_a = \frac{[OH⁻][H⁺]}{[H₂A]} \]
\[ K_c = \frac{[HA²⁻]}{[H⁺]} \]
\[ [H⁺] = [A³⁻] + 2[HA²⁻] + 3[A²⁻]
\]

(2) Choosing Master Variables. Two species, called CX and CY(I), are chosen, and initial guesses are made for their concentrations. The basis for the initial guesses will be discussed later, but guesses generally are not critical and are values known to be smaller than the actual amount. In this problem the following guesses were used:

\[ CX = 10^{-3} \text{(initial guess)} \]
\[ CY(I) = 10^{-4} \text{(initial guess)} \]

(3) Computing Other Variables. All but two of the equations are used, in a serial fashion, to calculate concentrations of remaining species. Simple algebraic combination of equations may sometimes be necessary.

\[ [H₂A] = \frac{[H⁺][A³⁻]}{K_a K_c} \]
\[ [HA²⁻] = \frac{[H⁺][A²⁻]}{K_c} \]
\[ [H₃A⁻] = K_c [H₂A] \]
\[ [OH⁻] = K_a \frac{[H⁺]}{[H₂A]} \]

Note that the sequence of calculations is important. Since the computer cannot operate in terms of unknown variables, [H₂A⁻] cannot be calculated in the third equation unless a value for [H₂A] has been calculated and stored in the computer.

(4) Formulation of Test Equations. The two remaining equations are now written in the following form:

\[ Y = [H₂A] + [H₃A⁻] + [HA²⁻] + [A³⁻] - C \]

and

\[ Y = [H₂A⁻] + 2[HA²⁻] + 3[A²⁻] + [OH⁻] - [H⁺] \]

and are referred to as test equations 1 and 2.

(5) Introduction of Problem Into Computer Program. The complete program for this problem written in FORTRAN is given in Figure 2. While it is beyond the scope of this discussion to delve into the details of programming, a brief explanation of the workings and statements of the program will be attempted. The essentials of FORTRAN, a widely used algebraic compiler designed for science and engineering problems, can be mastered in a few hours; a good introductory treatise is that of McCracken (4).

The coding of variables and constants is as follows:

\[ \begin{align*}
  [H⁺] &= CX \\
  [A³⁻] &= CY(I) \\
  [K₁] &= AK₁ \\
  [K₂] &= AK₂ \\
  [K₃] &= AK₃ \\
  [K_c] &= CH₃A \\
  [K_{OH⁻}] &= K_{OH⁻} \\
  [CH₃A] &= COH
\end{align*} \]

The concentrations of the species are all terms starting with the letter "C". The equilibrium constants use the letter "K" as a second letter but must begin with a letter other than I, J, K, L, M, or N, which are reserved for fixed point variables or integers in FORTRAN.

The enclosed portions of the program in Table 1 represent those parts added to the general program concerning the particular problem under consideration.

Part 1. Reading of equilibrium constants and analytical concentrations.

In this case the data given is for a solution in which \( C = 0.2 \), \( K₁ = 10^{-3} \), \( K₂ = 10^{-3} \), and \( K₃ = 10^{-13} \). FORTRAN exponential notation is used here, so that 1.0E-3 represents \( 1.0 \times 10^{-3} \).

![Table 1](image-url)

Figure 2. General computer program with statements for Example 1.
errors occur or wrong input data is used. If it is, the program leaves
and goes on to calculate CY(2). These steps are included to prevent wasting
time on the computer if errors occur or wrong input data is used.

The program now calculates a value of CY(1) which
satisfies test equation one. This is accomplished as
follows. A value of Y is calculated in statement 1 based on the initially guessed value of CY and the other
concentrations. Since the initial value of CY is guessed
smaller than the actual value, the sign of the value of
Y just calculated serves to indicate when subsequently
calculated CY-values are too small or too large. Therefore
this first value of Y is stored as a controlling number Z1 in statement 4. The parameter MY and statement 3 assure that Z1 is calculated only for the initially
guessed CY. Let us assume that Y is negative when
CY is too small. Then a negative value of Z1 is stored to
control the program. Whenever a tested value of CY appears which yields a negative value of Y, the
program will cause CY to increase. Whenever a positive
value of Y appears, the program will cause CY to
decrease. The correct value of CY will be that value
for which Y is zero. Y is tested by statements 6 or 8,
depending upon the sign of Z1. If CY(1) is too small, as it must be on the first trial, it is increased by 0.1 in
statement 9 (LY is initially equal to one) and the pro-
gram cycles back to statement 11 to try this new
CY(1). As long as CY(1) is too small it is increased in
steps of 0.1. When Y changes sign, indicating CY(1)
is now too large, CY(1) is decreased by 0.1 in statement 10, and increased by 0.01 in statement 9 (LY has
changed to two following statement 10). This process
continues until essentially ten significant figures have
been accumulated for CY(1), as tested by the IF-statement
two statements after statement 10.

Now CY(2) is treated in an identical manner using
test equation 2. The change from CY(1) to CY(2) is
accomplished using statements 12 and 13. After a value of CY(2) which satisfies test equation 2 (i.e.,
minimizes that value of Y) is calculated, the difference
between CY(1) and CY(2) is calculated in statement 14.
Now CX is varied using statements 24 on by a pro-
cedure virtually identical to that used for changing CY.
For each new value of CX two values of CY must be
calculated using the above procedure. At last, when a
value of CX which minimizes the value of Y in state-
ment 14 has been calculated, all of the equations are
simultaneously satisfied, and the answers are printed.

To illustrate how the program operated in this prob-
lem, values of CY(1), CY(2) and Y (in statement 14) for
every CX were printed and these results are shown in
Table 1. For the initial CX (10^-9) CY(1) and CY(2)
were calculated (the individual CY-values used in
calculating CY(1) and CY(2) are not given) and Y was
negative. CX was increased to 0.1, etc. Note that
CX is increased until Y changes sign. Although these
values are calculated to ten significant figures, only the
first four are shown. The computer program and
technique for varying CX and CY is similar to that used
by Herman and Bard for solving one-variable equations
(5). Although this method appears to be time-con-
suming, the total compiling and computation time for this
problem on a Control Data Corporation 1604
computer was only 17 seconds and the following results
were obtained:

\[
\begin{align*}
[H^+] &= 1.365 \times 10^{-3} M \\
[A^2-] &= 7.325 \times 10^{-3} M \\
[HA^-] &= 1.0 \times 10^{-6} M \\
[H_2A] &= 0.1863 M \\
[OH^-] &= 7.325 \times 10^{-3} M
\end{align*}
\]
Discussion

The variables selected to be represented as CX and CY(I) are generally those which appear most often in the equations. Thus, one generally selects [H\(^+\)] and complexing ligands since these repeatedly appear in equilibrium expressions. To solve directly for the remaining variables it may be necessary to combine some equations, although this is of minor difficulty.

The selection of the initial CX and CY(I) can be illustrated with Figure 3. CY(2) and CY(I) are plotted against CX. The intersection represents the correct values of CY(I) and CX. To the left of this point Y is negative, to the right it is positive. The program is written so that scanning is done in the increasing direction. The initial CX therefore must be smaller than the correct value and the initial CY must fall below both lines (such as point one). An initial CX close to the point of intersection allows the initial CY(I) to be guessed with a greater probability that it will fall below both lines. For example if CX is guessed too small, it is possible that the initial CY chosen will be above one of the CY lines at that point.

Another difficulty arises in the inability of the computer to retain more than eight or nine significant figures. In writing the sequence of equations one must be certain that a very low concentration isn't calculated as the difference between two large numbers. Thus, it would be poor to use the following equation:

\[ C = [H_2A] + [H_3A^-] + [HA^+] + [A^{2-}] \]

(5)

to calculate \([A^{2-}]\). This equation should be used either for calculating larger concentrations or as a test equation (as was done here).

![Figure 3. Plot of CY(I) versus CX.](image)

**Example 2**

Consider a solution which is 0.01 M in Ni\(^{2+}\), 0.1 M in NH\(_3\), 0.1 M in NH\(_4^+\), and C M in NaH\(_2\)Y, where H\(_2\)Y is ethylenediamine-tetraacetic acid (EDTA). This problem involves 17 solution species, including six nickel-ammonia complexes, the nickel-EDTA complex, and five ionization states of H\(_2\)Y. The various equilibrium constants are given below (5).

**Equilibrium Constants (Part 1):**

- Formation constants for Ni(NH\(_3\))\(^{2+}\):
  - \( K_1 = [\text{Ni(NH}_3\text{)}^{2+}] / [\text{Ni}^{2+}]\text{[NH}_3\text{]} = 562 \)
  - \( K_2 = 150.0 \)
  - \( K_3 = 48.0 \)
  - \( K_4 = 14.1 \)
  - \( K_5 = 5.12 \)
  - \( K_6 = 0.98 \)

**Ionization Constants for H\(_2\)Y:**

- \( K_{a1} = 1.0 \times 10^{-1} \)
- \( K_{a2} = 2.14 \times 10^{-2} \)
- \( K_{a3} = 6.97 \times 10^{-7} \)
- \( K_{a4} = 5.50 \times 10^{-11} \)
- \( K_{a5} = 1.16 \times 10^{-13} \)

**Variables and Initial Values:**

- Initial values for calculating larger concentrations or as a test equation (as was done here).

**Sequence of Equations (Part 2):**

\[ [\text{Ni(NH}_3\text{)}^{3+}] = K_1[\text{Ni(NH}_3\text{)}^{2+}]\text{[NH}_3\text{]} \]
\[ [\text{Ni(NH}_3\text{)}^{2+}] = K_2[\text{Ni(NH}_3\text{)}^{3+}]\text{[NH}_3\text{]} \]
\[ [\text{Ni(NH}_3\text{)}^{2+}] = K_3[\text{Ni(NH}_3\text{)}^{2+}]\text{[NH}_3\text{]} \]
\[ [\text{Ni(NH}_3\text{)}^{2+}] = K_4[\text{Ni(NH}_3\text{)}^{2+}]\text{[NH}_3\text{]} \]
\[ [\text{Ni(NH}_3\text{)}^{2+}] = K_5[\text{Ni(NH}_3\text{)}^{2+}]\text{[NH}_3\text{]} \]
\[ \text{[Ni}^{2+}] = 0.01 - [\text{Ni}^{2+}] - [\text{Ni(NH}_3\text{)}^{3+}] - [\text{Ni(NH}_3\text{)}^{2+}] - [\text{Ni(NH}_3\text{)}^{2+}] - [\text{Ni(NH}_3\text{)}^{2+}] \]
\[ [\text{NH}_4^+] = 0.2 - [\text{NH}_3] - [\text{Ni(NH}_3\text{)}^{3+}] - 2[\text{Ni(NH}_3\text{)}^{2+}] - 3[\text{Ni(NH}_3\text{)}^{2+}] - 4[\text{Ni(NH}_3\text{)}^{2+}] - 5[\text{Ni(NH}_3\text{)}^{2+}] - 6[\text{Ni(NH}_3\text{)}^{2+}] \]
\[ [\text{OH}^-] = K_g[\text{Ni}^{2+}] \]
\[ [\text{H}^+] = K_e[\text{OH}^-] \]
\[ [\text{Y}^-] = K_{a4}[\text{Y}^-] \]
\[ [\text{HY}^-] = [\text{H}^+][\text{Y}^-] \]
\[ [\text{H}_2\text{Y}^-] = [\text{H}^+][\text{HY}^-] \]
\[ [\text{H}_3\text{Y}^-] = [\text{H}^+][\text{H}_2\text{Y}^-] \]
\[ [\text{H}_4\text{Y}^-] = [\text{H}^+][\text{H}_3\text{Y}^-] \]
\[ [\text{H}_5\text{Y}^-] = [\text{H}^+][\text{H}_4\text{Y}^-] \]

**Test Equations:**

\[ Y = C - [\text{Ni}^{2+}] - [\text{H}_2\text{Y}^-] - [\text{H}_3\text{Y}^-] - [\text{H}_4\text{Y}^-] - [\text{H}_5\text{Y}^-] - [\text{Ni}^{2+}] - [\text{Ni(NH}_3\text{)}^{2+}] - [\text{Ni(NH}_3\text{)}^{2+}] - [\text{Ni(NH}_3\text{)}^{2+}] - [\text{Ni(NH}_3\text{)}^{2+}] - [\text{Ni(NH}_3\text{)}^{2+}] \]

**Table 2. Results for Example 2**

<table>
<thead>
<tr>
<th>C = 0</th>
<th>C = 0.005</th>
<th>C = 0.01</th>
<th>C = 0.012</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Ni(^{3+})] = 3.17 \times 10^{-10}</td>
<td>1.11 \times 10^{-10}</td>
<td>2.44 \times 10^{-10}</td>
<td>2.92 \times 10^{-10}</td>
</tr>
<tr>
<td>[Ni(^{3+})] = 1.16 \times 10^{-10}</td>
<td>4.51 \times 10^{-10}</td>
<td>1.10 \times 10^{-10}</td>
<td>8.85 \times 10^{-10}</td>
</tr>
<tr>
<td>[Ni(^{3+})] = 1.21 \times 10^{-10}</td>
<td>5.18 \times 10^{-10}</td>
<td>1.40 \times 10^{-10}</td>
<td>1.09 \times 10^{-10}</td>
</tr>
<tr>
<td>[Ni(^{3+})] = 3.98 \times 10^{-10}</td>
<td>1.83 \times 10^{-10}</td>
<td>5.23 \times 10^{-10}</td>
<td>4.18 \times 10^{-10}</td>
</tr>
<tr>
<td>[Ni(^{3+})] = 5.55 \times 10^{-10}</td>
<td>1.86 \times 10^{-10}</td>
<td>6.31 \times 10^{-10}</td>
<td>4.59 \times 10^{-10}</td>
</tr>
<tr>
<td>[Ni(^{3+})] = 1.19 \times 10^{-10}</td>
<td>6.90 \times 10^{-10}</td>
<td>2.65 \times 10^{-10}</td>
<td>1.83 \times 10^{-10}</td>
</tr>
<tr>
<td>[Ni(^{3+})] = 7.60 \times 10^{-10}</td>
<td>4.88 \times 10^{-10}</td>
<td>1.96 \times 10^{-10}</td>
<td>1.40 \times 10^{-10}</td>
</tr>
<tr>
<td>[NH(_3)] = 0.060</td>
<td>0.010</td>
<td>0.120</td>
<td>0.122</td>
</tr>
<tr>
<td>[H(_2)Y^-] = 0.100</td>
<td>0.110</td>
<td>0.120</td>
<td>0.122</td>
</tr>
<tr>
<td>[H(_3)Y^-] = 0.060</td>
<td>0.0722</td>
<td>0.0722</td>
<td>0.0722</td>
</tr>
</tbody>
</table>

*Time... 80 seconds.*
program can be used conveniently to calculate concentrations for many different conditions by using DO-loops for generating data (e.g., C in the case above). For example the calculation in Table 2 could represent four points in the titration of nickel(II) with EDTA in an ammoniacal medium. Similarly, to calculate the concentrations of the various species during the titration of H₃A with sodium hydroxide only the second test equation need be changed to

\[ Y = [H_3A^-] + 2[HA^{2-}] + 3[A^{3-}] + [OH^-] - [H^+] - B \]

where \( B \) represents the moles of sodium hydroxide added per liter. A DO-loop for generating various values of \( B \) would then calculate the titration curve of H₃A with sodium hydroxide. Library programs are available (7) to allow automatic plotting of the titration curve by the computer.

The program written here, or modifications of it, should prove useful for solving a large number of different kinds of problems in chemical equilibrium and for certain types of simultaneous, non-linear algebraic equations. The program presented must not be used without some background in computer programming.

For example, there are some variations in versions of FORTRAN for different computers which would require modification of the program presented. Appendix 1 in McCracken (4) might be useful for rewriting this program in another version of FORTRAN.

**Literature Cited**


(7) For example, UTM-2-02-033, "Point plot tape generator for CDC 1604 digital computer to the CDC 165 digital plotter," The University of Texas Computation Center, Austin, Texas (1962).