Calculation of optimum multicomponent solvent compositions for polymers using a multiple regression program

D. W. Behnken
American Cyanamid Company, Stamford, Connecticut, USA
(Received 26 January 1982; revised 19 April 1982)

A useful prediction of the solubility of a polymer in a particular solvent can be obtained by comparing their respective solubility parameters. The solution to the problem of finding solvent mixtures with maximum polymer interaction, already available for binary and ternary mixtures, has been extended to cover any number of components. This report also points out that the theoretical optimum composition (for two or more components) can be calculated by using a multiple regression program. The easy availability of such a program on most computers offers the occasional user an alternative to writing a special program for the purpose.

Keywords Solvents; polymer solvents; multicomponent solvents; solubility parameters; regression; optimum mixtures

INTRODUCTION

In recent papers Froehling *et al.*[1,2] and Rigbi[3] have discussed the problem of predicting optimum solvent mixtures for polymers and have provided solutions for binary[3] and ternary[2] mixtures. This work is based on the solubility parameter model (Hildebrand[4]) which states that the energy of mixing is related to the energies of vaporization of the pure components by the expression $\Delta E_{mix} = \lambda_1 \lambda_2 V(\delta_1 - \delta_2)^2$, where ΔE_{mix} is the energy of mixing, λ_1 and λ_2 are the volume fractions of the components, V is the volume of a mole of the mixture and δ_1 and δ_2 are the mixture component solubility parameters expressed as the square root of the cohesive energy densities. Letting ΔE_{vap} be the energy of vaporization and V_m the molar volume, $\delta^2 = \Delta E_{vap}/V_m$.

Using Hansen[5], δ^2 can be described as the sum of three contributing components, $\delta^2 = \delta_d^2$ and $\delta_p^2 + \delta_h^2$ where δ_d, δ_p and δ_h are the respective solubility parameter contributions from dispersion forces, dipole forces and hydrogen bonding (or donor–acceptor) interactions. Since the solubility parameters of a mixture are simply the (volume) average of the solubility parameters of the components, the problem of finding a good solvent mixture is that of finding the volume fractions that produce the closest possible solubility parameter match to the polymer, thereby minimizing the energy of mixing. The concept of closeness, however, involves distance in the three-dimensional space defined by the solubility parameter components δ_d, δ_p, δ_h and *not* simply the magnitudes of the respective δ's.

It must be mentioned that interactions not predicted by this simple theory can occur (particularly strong hydrogen bonds) which will spoil the predictive ability of the method. In general, however, the approach will provide useful estimates of solubility.

BINARY AND TERNARY SOLVENTS

Using the notation of Froehling and Hillegers[1] the solubility parameters (for dispersion, polar and hydrogen bonding forces) are represented by three-dimensional vectors. For solvents we use the vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \ldots$, for the polymer, $\boldsymbol{p}$, and for the solubility parameters of the optimum solvent mixture, the vector $\boldsymbol{m}$. Departing slightly from their notation, we let $\lambda_1, \lambda_2, \lambda_3, \ldots$ represent the volume fractions of the solvents $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \ldots$ respectively, in the optimum mixtures. As the references[1,2] point out, the optimum solution, $\boldsymbol{m}$, is obtained by finding the perpendicular projection of $\boldsymbol{p}$ onto the line between the endpoints of $\boldsymbol{a}$ and $\boldsymbol{b}$ for a binary mixture or onto the plane through the $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ endpoints for a ternary solvent mixture. This provides a solvent mixture whose solubility parameters are closest to those of the polymer as measured by the distance (Δ) in the three-dimensional solubility parameter space.

For a binary solvent this implies that λ_2, the optimum volume fraction of $\boldsymbol{b}$ in the mixture, is obtained by solving the equation

$$(\boldsymbol{b}-\boldsymbol{a})'(\boldsymbol{b}-\boldsymbol{a})\lambda_2 = (\boldsymbol{b}-\boldsymbol{a})'(\boldsymbol{p}-\boldsymbol{a}) \tag{1}$$

or

$$\lambda_2 = \frac{\sum_{j=1}^{3}(b_j - a_j)(p_j - a_j)}{\sum_{j=1}^{3}(b_j - a_j)^2} \tag{2}$$

where the subscripts ($j = 1,2,3$) denote the three elements of the corresponding vectors. The proportion of $\boldsymbol{a}$ is obtained from

$$\lambda_1 = 1 - \lambda_2$$

0032-3861/83/020196-06$03.00

and the solubility parameters of the optimum mixture from

$$\boldsymbol{m}=\lambda_1\boldsymbol{a}+\lambda_2\boldsymbol{b} \tag{3}$$

In multiple regression analysis[6], data are used to fit an equation of the general form

$$y=b_0+b_1x_1+b_2x_2+\ldots+b_kx_k \tag{4}$$

expressing a dependent variable, y, as a function of k independent variables x_1, x_2 ..., x_k. The intercept b_0 is optional and may be excluded from the model if desired. A set of N observations on each of the variables is needed as input data to fit the model and these data can be represented by the N-dimensional vectors $\boldsymbol{y}, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$. The program provides least-squares estimates of the regression coefficients $b_0, b_1, \ldots, b_k$ as well as other useful calculations to be described later.

To adapt this to the solvent problem we let the number of observations correspond to the dimensions of the solubility parameter vectors, so in all cases $N=3$. The intercept β_0 is deleted from the model. Then for a binary solvent we fit the equation

$$p=\lambda_1a+\lambda_2b \tag{5}$$

subject to the restriction that $\lambda_1+\lambda_2=1$. The easiest way to contend with this restriction is to substitute $\lambda_1=1-\lambda_2$ and restate the model as

$$p-a=\lambda_2(b-a) \tag{6}$$

Using the solubility parameter data to calculate the necessary input for the regression program let

$$\begin{aligned}\boldsymbol{y}&=\boldsymbol{p}-\boldsymbol{a}\\ \boldsymbol{x}_1&=\boldsymbol{b}-\boldsymbol{a}\end{aligned} \tag{7}$$

and fit the regression model

$$y=b_1x_1 \tag{8}$$

The estimated regression coefficient, b_1, provided by the program is identical to the value obtained from equation (2). Therefore

$$\begin{aligned}\lambda_2&=b_1\\ \lambda_1&=1-\lambda_2\end{aligned}$$

For a ternary solvent the same approach yields the regression input variables

$$\begin{aligned}\boldsymbol{y}&=\boldsymbol{p}-\boldsymbol{a}\\ \boldsymbol{x}_1&=\boldsymbol{b}-\boldsymbol{a}\\ \boldsymbol{x}_2&=\boldsymbol{c}-\boldsymbol{a}\end{aligned} \tag{9}$$

needed to fit the regression equation

$$y=b_1x_1+b_2x_2 \tag{10}$$

The regression coefficient estimates provided by the multiple regression program are identical to those obtained from equation (4) in Froehling and Hillegers[1]. The volume fractions λ_1, λ_2 and λ_3 for components $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ respectively are obtained from

$$\begin{aligned}\lambda_1&=1-b_1-b_2\\ \lambda_2&=b_1\\ \lambda_3&=b_2\end{aligned} \tag{11}$$

The other values of interest to those calculating optimum solvent compositions are also available from a typical multiple regression program. The three predicted values $\hat{y}_1$, $\hat{y}_2$ and $\hat{y}_3$, obtained by substituting the estimated regression coefficients into equations (8) or (10) at the three input values of x_1 (and x_2) form a vector $\hat{\boldsymbol{y}}$. These values are provided by the program and can be used to calculate the solubility parameters of the optimum mixture, namely

$$\boldsymbol{m}=\hat{\boldsymbol{y}}+\boldsymbol{a} \tag{12}$$

Alternatively, they can be obtained directly from $\boldsymbol{m}=\lambda_1\boldsymbol{a}+\lambda_2\boldsymbol{b}$ binary solvent or $\boldsymbol{m}=\lambda_1\boldsymbol{a}+\lambda_2\boldsymbol{b}+\lambda_3\boldsymbol{c}$ ternary solvent or from the residuals which are also printed out by most regression programs. The three residuals, defined as the difference between the input $\boldsymbol{y}$ values and the predicted values $\hat{\boldsymbol{y}}$ can be though of as a vector, $\boldsymbol{r}$

$$\boldsymbol{r}=\boldsymbol{y}-\hat{\boldsymbol{y}} \tag{13}$$

Using these residuals, the optimum mixture parameters can be calculated from the expression

$$\boldsymbol{m}=\boldsymbol{p}-\boldsymbol{r} \tag{14}$$

The distance, Δ, of the optimum solvent mixture from the polymer in the solubility parameter space is also available from the regression output. The residual sum of squares variously labelled as *RSS* or Error *SS* is defined as

$$RSS=\sum_{j=1}^{3}(y_j-\hat{y}_j)^2=\sum_{j=1}^{3}r_j^2 \tag{15}$$

and is equal to Δ^2. Therefore

$$\Delta=\sqrt{(\text{Residual sum of squares})} \tag{16}$$

The reason that the solution for the optimum solvent mixture is mathematically identical to the least-squares solution is that minimization of the distance Δ is in fact minimizing the residual sum of squares when the substitutions indicated by equations (7) and (9) are made. The equations given in refs. 1 and 2 for the λ_i are recognizable as the familiar normal equations encountered in least squares[6] when the restriction that $\Sigma b_i = 1$ is imposed.

If a flexible statistical package such as SAS (Statistical Analysis System) or Minitab is used, only the solubility parameter data need be entered, i.e. p, a, b, c. The other calculations such as equations (7), (9), (12), (14), (15) can be programmed for the computer to handle. Otherwise, a minor amount of hand calculation is required.

Example: Ternary solvent mixture (solubility parameter data from Froehling and Hillegers[1])

Solubility parameters $(J^{1/2}\,cm^{-3/2})$	p Nylon-6,6	a Water	b Trichloro-ethylene	c *n*-Propanol
δ_d	18.54	12.28	17.97	14.90
δ_p	5.12	81.30	3.10	6.70
δ_h	12.28	34.20	5.30	17.40

The regression variables are calculated from these to obtain:

$y = p - a$	$x_1 = b - a$	$x_2 = c - a$
6.26	5.69	2.62
−76.18	−78.20	−74.60
−21.92	−28.90	−16.80

The output obtained from a regression package would then supply the required values $b_1, b_2, \hat{y}, r$ and the residual sum of squares. From this we have the values given in Froehling and Hillegers[1]:

$$\lambda_2 = b_1 = 0.474$$

$$\lambda_3 = b_2 = 0.524$$

$$\lambda_1 = 1 - \lambda_2 - \lambda_3 = 0.002$$

The optimum mixture vector is given by

y	+	a	=	m
4.068		12.28		16.348
−76.128		81.30		5.172
−22.492		34.20		11.708

and $\Delta = \sqrt{(5.134\,428)} = 2.266$.

The regression approach can also be used when the object is to set one (or more) components at a fixed fraction of the solvent mixture and then find the optimum fractions for the other two. If we want a ternary mixture with the third component c at a specified fraction, $\mathring{\lambda}_3$, the procedure requires calculation of

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{b} - \boldsymbol{a} - \mathring{\lambda}_3(\boldsymbol{c} - \boldsymbol{a}) \\ \boldsymbol{x}_1 &= \boldsymbol{b} - \boldsymbol{a} \end{aligned} \tag{17}$$

and then fitting the model

$$y = b_1 x_1 \tag{18}$$

The volume fractions

$$\begin{aligned} \lambda_2 &= b_1 \\ \lambda_1 &= 1 - \lambda_2 - \mathring{\lambda}_3 \end{aligned} \tag{19}$$

and

$$\Delta = \sqrt{(\text{residual sum of squares})}$$

To obtain the optimum mixture parameters we can use either

$$\boldsymbol{m} = \boldsymbol{y} + \boldsymbol{a} + \mathring{\lambda}_3(\boldsymbol{c} - \boldsymbol{a}) \tag{20}$$

or the alternative approach for calculating $\boldsymbol{m}$ given by equation (14) which would be simpler in this case.

If we want the best solvent for nylon-6,6 in the example above, *given* that the mixture is 5% water, we need the following input calculations:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{p} - \boldsymbol{b} - 0.05(\boldsymbol{a} - \boldsymbol{b}) \\ \boldsymbol{x}_1 &= (\boldsymbol{c} - \boldsymbol{b}) \end{aligned} \tag{21}$$

Note here that the role of $\boldsymbol{a}$ was altered since water was the component selected to be fixed. The roles of $\boldsymbol{b}$ and $\boldsymbol{c}$ could of course be reversed except that b_1 would provide λ_2 instead of λ_3. The regression input would follow as:

$y = p - b - 0.05\,(a - b)$	$x_1 = (c - b)$
0.8545	−3.07
−1.8900	3.60
5.5350	12.10

After running the regression we have in hand the estimate of $b_1 = 0.341$, the residual sum of squares $= 15.3196$ and the vector of residuals $\boldsymbol{r}$.

From these we have:

$$\lambda_3 = 0.341$$

$$\lambda_2 = 0.050$$

$$\lambda_1 = 0.609$$

and

$$\Delta = \sqrt{(15.3196)} = 3.91$$

To calculate $\boldsymbol{m}$ directly we can either use equation (14), i.e.

p	−	r	=	m
18.54		1.901		16.348
5.12		3.117		5.172
12.28		1.410		11.708

or equation (20). In this case, however, $\boldsymbol{m} = \boldsymbol{y} + \boldsymbol{b} + 0.05\,(\boldsymbol{a} - \boldsymbol{b})$ (since $\boldsymbol{b}$, not $\boldsymbol{a}$, was used as the origin in this example).

Non-feasible solutions

In the examples given above, feasible solutions were obtained in that all the λ_i fell in the interval between 0 and 1 and the results could be interpreted as determining a mixture of components. The equations used to solve for the λ_i do not guarantee this. The problem they solve is to

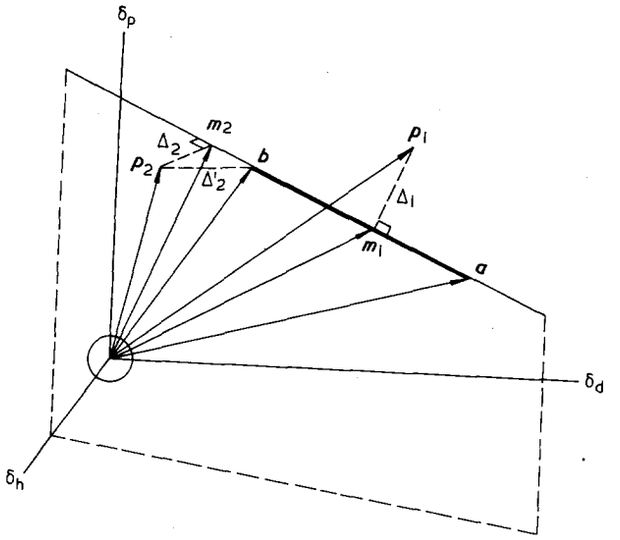

Figure 1 Binary solvent geometry

find the linear combination of the solvent vectors in solubility parameter space that minimizes the distance from the polymer vector, subject to the restriction that $\Sigma\lambda_i = 1$. They do *not* add the restriction

$$0 \leqslant \lambda_i \leqslant 1 \qquad i = 1,2,\ldots$$

As pointed out by Froehling and Hillegers[1] when negative values and values in excess of unity are obtained, it is simply saying that the best attainable solvent mixture is to be found with fewer components. Geometrically, for a binary solvent, the projection of $\boldsymbol{p}$ onto the line passing through the endpoints of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ fails to fall in the interval between them. This segment between $\boldsymbol{a}$ and $\boldsymbol{b}$ is the locus of all feasible solutions. In *Figure 1*, polymer $\boldsymbol{p}_1$ yields a solution with physical meaning ($0 \leqslant \lambda_2 \leqslant 1$) while $\boldsymbol{p}_2$ does not ($\lambda_2 > 1$). The best choice for this polymer is solvent $\boldsymbol{b}$ since the shortest distance to a feasible solution (Δ'_2) is obtained at $\boldsymbol{b}$ ($\lambda_2 = 1$, $\lambda_1 = 0$).

In *Figure 2a* the ternary solvent locus of feasible solutions (all $0 < \lambda_i \leqslant 1$ and $\sum\lambda_i = 1$) is represented by the shaded triangular region of the plane through the endpoints of $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$. The polymer $\boldsymbol{p}_1$ yields a solution that falls in this area. To visualize the projection problem more easily, an alternative picture of the space is shown in *Figure 2b*. The origin has been shifted to $\boldsymbol{a}$ by subtracting $\boldsymbol{a}$ from every vector, and the plane through $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ has been rotated into the position of the horizontal plane. All relative positions remain unchanged. For polymer $\boldsymbol{p}_2$, when one solves the equations and finds $\boldsymbol{m}_2$, the projection of $\boldsymbol{p}_2$ onto the plane of $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$, it falls outside the shaded area of feasible solutions. The shortest distance to a feasible solution is found by projecting onto the line connecting $\boldsymbol{a}$ and $\boldsymbol{c}$ which says that the optimum solvent $\boldsymbol{m}'_2$ is a mixture of solvents $\boldsymbol{a}$ and $\boldsymbol{c}$. Δ'_2 is the length of this shortest projection. A polymer such as $\boldsymbol{p}_3$ would find the shortest distance to the region of feasible solutions at $\boldsymbol{b}$ which says the pure solvent $\boldsymbol{b}$ provides the best match. (If the polymer vector $\boldsymbol{p}$ falls within the region of feasible solutions, the solvent match is perfect and $\Delta = 0$.)

This geometric insight illustrates why non-feasible solutions are obtained and shows that for these cases the best solution must lie on the boundary of the region. To find the best mixture then, once a non-feasible solution is obtained, the approach is to solve the problems formed by dropping each solvent one at a time and finally to calculate the pure solvent distances, e.g. for $\boldsymbol{a}$

$$\Delta = \sqrt{[(\boldsymbol{p}-\boldsymbol{a})'(\boldsymbol{p}-\boldsymbol{a})]} = \sqrt{\left(\sum_{i=1}^{3}(p_i - a_i)^2\right)} \qquad (22)$$

The latter is necessary since, even though an $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ mixture does not provide a feasible solution, and the two-component $\boldsymbol{b}, \boldsymbol{c}$ mixture *does* provide a useful result, when the Δ's are compared, the pure solvent $\boldsymbol{a}$ might provide the smaller value.

Figures 2a and *b* also illustrate the connection between the regression problem and the problem considered here. When the origin is shifted to point $\boldsymbol{a}$ by subtracting $\boldsymbol{a}$ from the other vectors and the labels $\boldsymbol{y}, \boldsymbol{x}_1, \boldsymbol{x}_2$ are substituted as indicated in equations (7) and (9), this is the geometry of the least-squares regression problem[6]. The restriction that $\sum\lambda_i = 1$ causes the shift of the origin, the choice of the particular solvent $\boldsymbol{a}$ being completely arbitrary. To find the least-squares solution the length Δ of the vector $\boldsymbol{r}$ (equations (13) and (16)) must be minimized, which results in the perpendicular projection onto the line, plane or hyperplane in which the $\boldsymbol{x}$ vectors lie. In a usual regression application of course, N, the number of elements in the vectors, is not limited to three.

FOUR- OR MORE COMPONENT SOLVENTS

For a four-component solvent the region of feasible solutions in general is bounded by a tetrahedron as shown

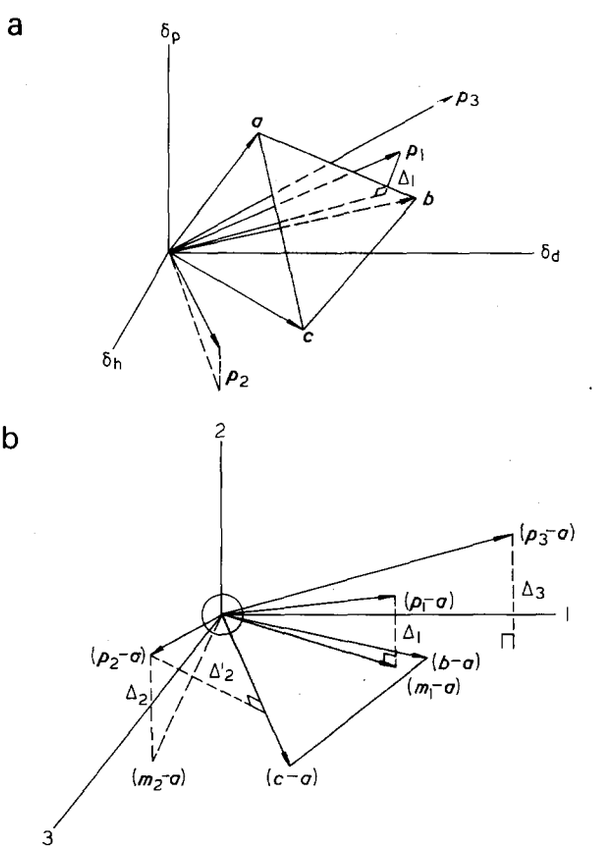

Figure 2 (a) Ternary solvent geometry. (b) Transformed ternary solvent geometry

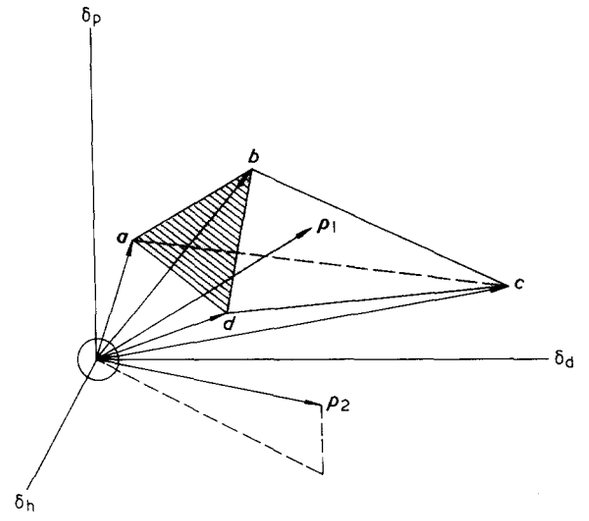

Figure 3 Four-component solvent geometry

in *Figure 3*. In this case, as well as for binary or ternary mixtures, dependencies among the solvent vectors are possible that reduce the dimensionality of this region. This implies that the composition of the optimum mixture is not unique, the same ***m*** vector being attainable by more than one (in fact an infinity of) composition(s). To illustrate geometrically, this occurs in a ternary mixture when the endpoints of ***a***, ***b*** and ***c*** all fall on the same line and in a four-component system when ***a***, ***b***, ***c*** and ***d*** endpoints fall in a plane. The interior solvent is redundant in these cases and can be eliminated from the problem if desired without changing the optimum vector ***m***.

In the general four-component case, regression using the variables

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{p} - \boldsymbol{a} \\ \boldsymbol{x}_1 &= \boldsymbol{b} - \boldsymbol{a} \\ \boldsymbol{x}_2 &= \boldsymbol{c} - \boldsymbol{a} \\ \boldsymbol{x}_3 &= \boldsymbol{d} - \boldsymbol{a} \end{aligned} \tag{23}$$

and the model

$$y = b_1 x_1 + b_2 x_2 + b_3 x_3 \tag{24}$$

will lead to a solution with $\Delta = 0$ as long as the vectors $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, and $\boldsymbol{x}_3$ are independent. In other words, the mixture will have the same solubility parameters as the polymer ($\boldsymbol{p} = \boldsymbol{m}$). Just as before, however, *only if* $\boldsymbol{p}$ lies within or on the surface of the feasibility region will the solution have any physical significance. If so, then

$$\begin{aligned} b_1 &= \lambda_2 \\ b_2 &= \lambda_3 \\ b_3 &= \lambda_4 \\ 1 - \lambda_2 - \lambda_3 - \lambda_4 &= \lambda_1 \end{aligned} \tag{25}$$

and all $0 \leqslant \lambda_i \leqslant 1$.

If the λ_is are not feasible, then all three-, two-, or one-component mixtures must be investigated for the smallest Δ as described earlier.

Example: Four-component solvent mixture (solubility parameter data from Froehling and Hillegers[1] and Rigbi[3])

Solubility parameters ($J^{1/2}$ cm$^{-3/2}$)	*p* Poly(methyl-methacrylate)	*a* n-Propanol	*b* Toluene	*c* MEK	**d** Chloro-form	$y = (p-a)$	$x_1 = (b-a)$	$x_2 = (c-a)$	$x_3 = (d-a)$
δ_d	15.74	14.9	18.05	15.7	17.7	0.84	3.15	0.8	2.8
δ_p	8.20	6.7	1.40	9.0	3.1	1.50	−5.30	2.3	−3.6
δ_h	6.70	17.4	2.00	5.1	5.7	−10.70	−15.40	−12.3	−11.7

The solution, resulting in a perfect match of the solubility parameters ($\Delta = 0$, $\boldsymbol{p} = \boldsymbol{m}$), for poly(methyl methacrylate) was found to be:

Solvent	Volume fraction	Regression coeff.
n-Propanol	$\lambda_1 = 0.140$	$1 - b_1 - b_2 - b_3$
Toluene	$\lambda_2 = 0.044$	b_1
MEK	$\lambda_3 = 0.792$	b_2
Chloroform	$\lambda_4 = 0.024$	b_3

It is mathematically possible that a polymer that fails to give feasible solutions with a set of four solvents can be brought inside the feasibility region by adding additional solvents which increase the size of the polyhedron. The computational approach then becomes an iterative one since the equations are of rank 3 and will only support solving for three unknowns, i.e. the solution is not unique. To do this using the regression approach would involve setting all but four of the solvents at fixed fractions, solving the resulting equations and checking the λ_is for feasibility. If not all $0 \leqslant \lambda_i \leqslant 1$ then a methodical scanning of a grid of possible values for $\lambda_5, \lambda_6, \ldots$ could be made, calculating trial solutions at each point.

For example, if six components ***a***, ***b***, ***c***, ***d***, ***e***, ***f***, were to be investigated, we might check a grid of values for λ_5 and λ_6 varying each in steps of 0.1 but keeping $(\lambda_5 + \lambda_6) < 0.9$. For a typical solution we might set $\lambda_5 = 0.1$ and $\lambda_6 = 0.2$, calculate the regression variables

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{p} - \boldsymbol{a} - \lambda_5(\boldsymbol{e} - \boldsymbol{a}) - \lambda_6(\boldsymbol{f} - \boldsymbol{a}) \\ \boldsymbol{x}_1 &= \boldsymbol{b} - \boldsymbol{a} \\ \boldsymbol{x}_3 &= \boldsymbol{c} - \boldsymbol{a} \\ \boldsymbol{x}_3 &= \boldsymbol{d} - \boldsymbol{a} \end{aligned} \tag{26}$$

and fit the model of equation (24) to find λ_1, λ_2, λ_3 and λ_4

($\lambda_1 = 1 - \sum_{i=2}^{6} \lambda_i$). Once a solution is found, an infinity of variations is possible in the neighbouring region of the solubility parameter space.

The regression approach would not be efficient unless a package such as SAS were available which could easily be instructed to generate the whole set of different y vectors and sequentially provide all the regression solutions. For values of λ_5 and λ_6 that produce solutions promisingly close to feasibility, it will be necessary to investigate this region with a finer grid.

Some results obtained using the above procedure with SAS to find optimum ($\Delta = 0$) six-component mixtures for poly(methyl methacrylate) are given below.

	Volume fractions		
Solvent	Mixture 1	Mixture 2	Mixture 3
Toluene	0.026	0.040	0.008
MEK	0.755	0.778	0.718
Chloroform	0.020	0.020	0.016
DMF	0.020	0.007	0.040
n-Propanol	0.120	0.134	0.100
Butyl acetate	0.059	0.021	0.118

NUMERICAL ASPECTS

In solving the linear equations to find the λ_i values, experience has shown that the equations may be very poorly conditioned as a result of close similarities in the solvent vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}, \ldots$ If this occurs, it is important that double-precision arithmetic is used on the computer to avoid obtaining grossly incorrect solutions. Hand calculation is not generally recommended. It should be noted that although the four- (or more) component problem can be solved by the regression approach, it is really the degenerate case of least squares in which the number of observations equals the number of unknowns. If the four-component problem were to be solved *directly*, the following set of three equations in three unknowns should be used instead of the least-squares normal equations

$$(\boldsymbol{b}-\boldsymbol{a})\lambda_2 + (\boldsymbol{c}-\boldsymbol{a})\lambda_3 + (\boldsymbol{d}-\boldsymbol{a})\lambda_4 = (\boldsymbol{p}-\boldsymbol{a}) \qquad (27)$$

The extension to more components simply involves subtraction of the fixed components from the right-hand side as suggested by equation (26). Some multiple regression programs may not have adequate checks built in to cope with cases that result in zero values for the residual sum of squares. No problems should be encountered in calculating the regression coefficients but error messages may result from the subsequent statistical analysis which involves division by functions of the residual sum of squares.

SUMMARY

A computer approach is described which utilizes any multiple regression package for calculating the composition of a solvent mixture that gives maximum interaction with a polymer. The only requirement is that the program permits fitting the general model

$$y = b_1 x_1 + b_2 x_2 + \ldots + b_k x_k$$

without a leading constant term (intercept). The input variables needed for the computer are easily calculated by hand from the polymer and solvent solubility parameter vectors. If a regression package such as SAS or Minitab is available, it can be easily programmed to perform these calculations so that only the solubility parameters need be entered.

The approach has been extended to cover mixtures of more than three components and to allow optimum mixtures of solvents to be formed when one or more components are preset at fixed levels.

ACKNOWLEDGEMENT

The assistance of P. M. D. Benoit in programming the SAS computations is gratefully acknowledged.

REFERENCES

1 Froehling, P. E. and Hillegers, L. T. *Polymer* 1981, **22**, 261
2 Froehling, P. E., Koenhin, D. M., Bantjes, A. and Smolders, C. M. *Polymer* 1976, **17**, 835
3 Rigbi, Z. *Polymer* 1978, **19**, 1229
4 Hildebrand, J. H. and Scott, R. L. 'The Solubility of Non-Electrolytes', 3rd Edn., Dover Publ. New York, 1949
5 Hansen, C. M. *J. Paint Technol.* 1967, **39**, 104
6 Draper, N. R. and Smith, H. 'Applied Regression Analysis', 2nd Edn., Wiley, New York, 1981