APPENDIX A - AN EXAMPLE DATA FILE

The file below was written out by COLETTE ready for input to FISH, it has been reduced to 6 data points, the third record has been edited by the user so that only points 2-4 are to be used in FISH. The 3 in the fith record shows that Q, I(Q) & error(Q) are expected, using the format shown.

```
LOQ Wed 18-MAY-1988 15:21  SAMPLE:   714     EMPTY CAN:   749
lambda  2.20 10.00   Phi  -90. >   90. deg   Radii  35.0 335.
    0   6    0    0    0    2    4
        0         0         0         0
 3 (F12.5,2E16.6)
     0.00562    1.664269E+01    1.182694E-01
     0.00607    1.018861E+01    6.170455E-01
     0.00655    4.091472E+00    3.789476E-01
     0.00707    4.746222E+00    4.646616E-01
     0.00865    6.092464E+00    2.959473E-01
     0.00947    8.743887E+00    2.343611E-01
```

APPENDIX B - AN EXAMPLE MODEL FILE - LSINP.DAT

This example file contains three separate models: (i) a single spherical particle plus background, note that only the constant term in the cubic background is being used to give a flat background addition.

(ii) polydisperse spheres with $R^{-1.5}$ size distribution defined by Rmid= 60 and width lambda =0.5. The lambda and R-MID parameters 5 & 6 do not do anything directly so are MODEL 2; they are related to R1 and R2 via the two constraints. Again a flat background is included.

(iii) similar polydisperse spheres interacting via a hard spheres structure factor. The hard sphere radius and volume fraction have partial shifts of 0.4 in order to damp their changes. Instrument resolution smearing is also added. The label field is entirely optional, the calculation is controlled by the MODEL and LTYP numbers in the second and thid columns.

As a reminder the main columns are (note the running count number is ignored on input):

```
Running Model LTYP(i) Label Parameter  ESD       Partial  Calc
count   number              Value V()           shift    shift
 i                                              On/Off
 I3     I3    I3       A12  E14.6      E13.3     F6.1      E10.2

 T  1 P  6 S  0 C  0 N  0

 W  1 K  0 IP 1 MS 1 IY 1 Q  7 R  7    6    1

   SINGLE SPHERE PLUS BKG

  1  1  1 C1           1.842795E-04    5.014E-06   1.0  4.96E-08

  2  1  2 R1           1.000000E 02    0.000E+00   1.0  0.00E+00

  3  3  1 BKG A        9.848869E-02    1.601E-02   1.0  2.20E-05

  4  3  2     B*Q      0.000000E+00    0.000E+00   0.0  0.00E+00

  5  3  3     C*Q**2   0.000000E+00    0.000E+00   0.0  0.00E+00

  6 99  1    SCALE     1.000000E+00    0.000E+00   0.0  0.00E+00

 T  1 P 10 S  1 C  2 N  5

 W  1 K  0 IP 1 MS 1 IY 1 Q  7 R  7    6    1

  POLY DISP SPHERES PLUS BKG    CONSTRAINED TO FIT RM AND LAM

  1  6 81 R**N VOL     1.842795E-04    5.014E-06   1.0  4.96E-08

  2  6  2 R2-R1        7.326303E+01    0.000E+00  -1.0  0.00E+00
```

```
  3  6  3 R1           2.997124E+01    0.000E+00  -1.0  0.00E+00
  4  6  4 N          -1.500000E+00    0.000E+00   0.0  0.00E+00
  5  2  1 LAMBDA       5.000000E-01    0.000E+00   0.0  0.00E+00
  6  2  2    R-MID     6.000000E+01    4.136E-01   1.0  2.46E-03
  7  3  1 BKG A        9.848869E-02    1.601E-02   1.0  2.20E-05
  8  3  2     B*Q      0.000000E+00    0.000E+00   0.0  0.00E+00
  9  3  3     C*Q**2   0.000000E+00    0.000E+00   0.0  0.00E+00
 10 99  1    SCALE     1.000000E+00    0.000E+00   0.0  0.00E+00
   1  1 2114.SUB     CALC 2 BKG 3 POL 4  SSE= 3.898E+01
   3   2   5   6   0
   0.00000   0.00000   0.00000   0.00000
   4   3   5   6   0
   0.00000   0.00000   0.00000   0.00000
 1.000E-01 1.843E-04 3.112E-01 2.000E+00 1.032E+02
T  1 P 14 S  1 C  2 N  5
 W  1 K  0 IP 1 MS 1 IY 1 Q  7 R  7    6    1
  (POLY DISP SPHERES)*HSS(Q)    CONSTRAINED TO FIT RM AND LAM
  1  6 81 R**N VOL     1.842795E-04    5.014E-06   1.0  4.96E-08
  2  6  2 R2-R1        7.326303E+01    0.000E+00  -1.0  0.00E+00
  3  6  3 R1           2.997124E+01    0.000E+00  -1.0  0.00E+00
  4  6  4 N          -1.500000E+00    0.000E+00   0.0  0.00E+00
  5  2  1 LAMBDA       5.500000E-01    0.000E+00   0.0  0.00E+00
  6  2  2    R-MID     6.660275E+01    4.136E-01   1.0  2.46E-03
  7 22  1 HS S(Q) VOL  1.854555E-01    1.444E-02   0.4 -4.55E-05
  8 22  2 SPH RADIUS   1.351837E+02    2.197E+00   0.4  2.26E-01
  9  3  1 BKG A        9.848869E-02    1.601E-02   1.0  2.20E-05
 10  3  2     B*Q      0.000000E+00    0.000E+00   0.0  0.00E+00
 11  3  3     C*Q**2   0.000000E+00    0.000E+00   0.0  0.00E+00
 12 15  1 SMEAR        1.000000E+00    0.000E+00   0.0  0.00E+00
 13 15  2 NSIMP        2.100000E+01    0.000E+00   0.0  0.00E+00
 14 99  1    SCALE     1.000000E+00    0.000E+00   0.0  0.00E+00
   1  1 2114.SUB     CALC 2 BKG 3 POL 4  SSE= 3.898E+01
   3   2   5   6   0
   0.00000   0.00000   0.00000   0.00000
   4   3   5   6   0
   0.00000   0.00000   0.00000   0.00000
 1.000E-01 1.843E-04 3.112E-01 2.000E+00 1.032E+02
```

APPENDIX C - SPECIAL MODELS FOR 2-SHELL ELLIPSOID

MODEL 24 with LTYP(1)=21 or 31 are for swollen micellar systems with either constant ratios of inner and outer radii or constant outer shell thickness respectively. The standard model is entirely reparametrised using most of the numerical constant CON() array to store physical parameters. [Note this precludes the use of polydisperse spheres or other models using the CON() array, though the CON() here could easily be moved if required.]

The model assumes a surfactant micelle in water where the surfactant has a hydrocarbon tail. The head group/water interface may be "staggered", thereby including some of the surfactant tails in the "head group" region. This is controlled by ALF. One terminal CH3 group is always assumed to be in the core (compare CON(2&3) with CON(15&16). Oil may be added to the core region via CON(22-24).

LTYP=21 or 31	SCALE	
2	AGG	aggregation number, number of surfactant molecules per micelle
3	X	axial ratio (if LTYP(1)=21) is not used if LTYP(1)=31 when shell thickness δ is calculated in CON(19) [Check this, I'm not sure !]
4	Charge	
5	ALF	fraction of CH2 chain groups in core (α)

SCALE = $(\rho_2 - \rho_3)^2 V_T^2 N$ where number density of micelles N = [conc-cmc]N_A/(agg. no.) and SCALE has units of cm^{-1}. The fitted value of this parameter should agree with the expected value in CON(17), if all the conversions to absolute units are correct !

This model makes extensive use of the numerical constants array, so it may be incompatible with some other models.

CON(1)=N1	Marquardt lambda, as usual
2	CH_3 volume, $Å^3$
3	Σbi CH_3, in units of 10^{-12}cm
4	head group volume, $Å^3$
5	Σbi headgroup, scattering length in units of 10^{-12}cm
6	hydration number headgroup
7	counterion volume, $Å^3$
8	Σbi counterion, in units of 10^{-12}cm
9	counterion hydration number
10	solvent volume, $Å^3$
11	Σbi solvent, in units of 10^{-12}cm
12	scattering length density of solvent, ρ_3, $Å^{-2}$
13	[conc-cmc] surfactant, mol/litre
14	NA, Avogadro's number.10^{-19}= 6.0235×10^4
15	$(CH2)_{m-1}$ volume, $Å^3$
16	Σbi " in units of 10^{-12}cm
---	---
17	expected scale in cm-1 17-21 are calculated by the program
18	R1, Å
19	R2 or δ

20 ρ_1, Å^{-2}

21 ρ_2

---

22 [oil], concentration of solubilised oil, mol/litre,
note N_{oil} per micelle = aggregation number.[oil]/[conc-cmc]

23 oil volume, Å3

24 Σbi oil, cm

These parameters are used by routine AGGRE to calculate the normal parameters to pass to the routines for two-shell ellipsoids ELLSH1 and ELLSH2.

Guides to use: ALWAYS check CON(17)-CON(21) to see if their values are reasonable.

CON(17) should agree (say within 20%) with the refining SCALE parameter, assuming the absolute units are correct.

CON(18), R1 or R1.X, should be less than the fully stretched tail, unless extra oil is present.

The calculation proceeds by calculating the volume of the inner core:

$$V_1 = (V_{CH3} + ALF.V_{CH2(m-1)}).AGG$$

Inner radius $R_1 = (3V_1/4\pi X)^{1/3}$

Total volume of a micelle VT = [Vchain + Vhead group + Vion +

Vsolvent.(hydration number of head group + hydration number of ion)].AGG

-

(Vion + Vsolvent.(hydration number ion)).CHARGE

The value of R2 (LTYP(1)=21) is given by $V_T = (4\pi/3)R_2^3X$

Shell thickness δ (LTYP(1)=31) is given by $V_T = (4\pi/3)(R_1+\delta)^2(R_1X + \delta)$ so δ is the only positive solution of:

$$\delta^3 + R(X+2)\delta^2 + R^2(2X+1)\delta - 3(V_T - V_1)/4\pi = 0$$

Knowing $V_1$and V_2 and the scattering lengths the program calculates ρ_1 and ρ_2 and the coefficient of $j_1(u_1)/u_1$. Adding oil makes the above equations a little more complicated.

APPENDIX D - AN EXAMPLE INTERACTIVE SESSION

Input typed by the user is underlined, text in [square brackets] are explanatory comments added later.

$@runfish

```
Welcome to the FISH data analysis program, enter HELP
         after  Command>   if you are lost.
         please type commands using only UPPER CASE
Command> READ 2658.Q
```

[start sequential reading of data file]

```
DATA INPUT ROUTINE, READS FROM FORTRAN STREAM NO. 1
UP TO 9 SETS ("WORKSPACES") MAY BE USED
ENTER 0 (RETURN) TO IGNORE, OVERWRITES PREVIOUS SET OF SAME NO.
91 TO REWIND,  99 TO STOP,  95 TO SEARCH
```

```
 LOQ Mon 22-MAY-1989 16:42SAMPLE:  2658     EMPTY CAN:  2646

 Lambda  2.00 10.00 Phi  -90. >   90. deg Radii  35.0 335.0

SAVE AS SET NO. = ? 1
```

[lets keep this data as set number 1]

```
 ENTER LOCAL LABEL FOR THIS SET  (A11)= 2658
3 (F12.5,2E16.6)
 LOQ Mon 22-MAY-1989 16:42SAMPLE:  2658     EMPTY CAN:  2646
Lambda  2.00 10.00 Phi  -90. >   90. deg Radii  35.0 335.0
SAVE AS SET NO. = ? 99
```

[get out of the input loop]

[lets try to fit the data ...]

```
Command> FIT
1-READ MODEL FILE              4-CALC ONLY, CHOOSE Q
2-CHOOSE OBS,CALC SETS ETC.    5-INDEX
3-ENTER FIT ROUTINE            6-RETURN TO MAIN ROUTINE
1
```

[start to read model file LSINP.DAT]

```
LSQIN READS MODEL DESCRIPTION FILE FROM FORTRAN STREAM 3
ENTER 0-TO IGNORE, 1-USE THIS MODEL, 9-REWIND
 Gaussian coil
 for polymers
 1 14  1 Const        3.500000E+02    0.000E+00   0.4  0.00E+00
 2 14  2 Rg           7.500000E+01    0.000E+00   0.4  0.00E+00
 3  3  1 BKG A        5.000000E+00    0.000E+00   0.4  0.00E+00
 4  3  2   B*Q        0.000000E+00    0.000E+00   0.0  0.00E+00
USE=
```

[prints first 4 record of each model in file]

[pressed return to enter zero, and read next model]

```
  SINGLE SPHERE PLUS BKG
 1  1  1 C1           1.842795E-04    5.014E-06   1.0  0.00E+00
 2  1  2 R1           1.000000E+02    0.000E+00   1.0  0.00E+00
 3  3  1 BKG A        9.848869E-02    1.601E-02   1.0  0.00E+00
 4  3  2     B*Q      0.000000E+00    0.000E+00   0.0  0.00E+00
USE=0

  POLY DISP SPHERES PLUS BKG    CONSTRAINED TO FIT RM AND LAM
 1  6 81 R**N VOL     1.842795E-04    5.014E-06   1.0  0.00E+00
 2  6  2 R2-R1        7.326303E+01    0.000E+00  -1.0  0.00E+00
```

```
 3   6   3 R1            2.997124E+01    0.000E+00  -1.0  0.00E+00
 4   6   4 N           -1.500000E+00    0.000E+00   0.0  0.00E+00
USE=1                [use this model and get out of input loop]

1-READ MODEL FILE              4-CALC ONLY, CHOOSE Q
2-CHOOSE OBS,CALC SETS ETC.    5-INDEX
3-ENTER FIT ROUTINE            6-RETURN TO MAIN ROUTINE
2                          [now proceed around the main fit menu ]
NO. OF SETS OF DATA TO FIT ( MAX 3, (I1) ) = 1

                            ENTER DATA SET numbers for:
                           OBS  observations i.e. data to be fitted.
                           CALC for results of model calculation.
                           BKG  if you are about to use experimental
                               subtraction, model 4, ( WORK is ALSO
required !)
                           POLY is needed for polydispersity e.g. for
model 6
                           WORK is for polynomial background, model 3,
                               and/or observed background model 4, which
are not
                               smeared so are kept separate.
                           UNSMEARED optional to save unsmeared when
smearing.
                           P(Q) optional to save form factor separately.
                           S(Q) optional to save structure factor
separately.
                           BETA(Q) optional beta ratio for polydisperse.
                           DEBUG(Q) optional for anything else !
                           Enter zeroes (spaces) for sets not required.

OBS, CALC, BKGD, POLY, WORK, UMSMEARED, P(Q), S(Q), BETA(Q), BUG(Q)
 ENTER SET NUMBERS FOR EACH (10I1)
1203456780                 [ decides which data set to fit and where to store the output ]

REFINING  1 SETS   OBS CAL BKG PLY WRK USM  PQ  SQ BET DBG
      2658          1   2   0   3   4   5   6   7   8   0
1-READ MODEL FILE              4-CALC ONLY, CHOOSE Q
2-CHOOSE OBS,CALC SETS ETC.    5-INDEX
3-ENTER FIT ROUTINE            6-RETURN TO MAIN ROUTINE
3
```

```
FITTING/CALCULATION ROUTINE, TYPE PP TO SEE THE MODEL,
RUN TO CALCULATE, HELP FOR A LIST OF COMMANDS
>
PP                    [look at whole model ]
T  1 P 10 S  1 C  2 N  5
W  1 K  0 IP 1 MS 1 IY 1 Q -6 R -6  -11  -11    1    0    2    0    0
0    0
  POLY DISP SPHERES PLUS BKG    CONSTRAINED TO FIT RM AND LAM
 1  6 81 R**N VOL     1.842795E-04    5.014E-06   1.0  0.00E+00
 2  6  2 R2-R1        7.326303E+01    0.000E+00  -1.0  0.00E+00
 3  6  3 R1           2.997124E+01    0.000E+00  -1.0  0.00E+00
 4  6  4 N           -1.500000E+00    0.000E+00   0.0  0.00E+00
 5  2  1 LAMBDA       5.000000E-01    0.000E+00   0.0  0.00E+00
 6  2  2    R-MID     6.000000E+01    4.136E-01   1.0  0.00E+00
 7  3  1 BKG A        9.848869E-02    1.601E-02   1.0  0.00E+00
 8  3  2     B*Q      0.000000E+00    0.000E+00   0.0  0.00E+00
 9  3  3     C*Q**2   0.000000E+00    0.000E+00   0.0  0.00E+00
10 99  1    SCALE     1.000000E+00    0.000E+00   0.0  0.00E+00
  1  1 2658         CALC 2 BKG 0 POL 3  SSE= 0.000E+00
CONSTRAINT  1 TYPE  3 USE= 1
V(  2 ) = 2.0* V(  5 ) * V(  6 )   R2-R1=2*L *RM

CONSTRAINT  2 TYPE  4 USE= 1
V(  3 ) = (1.0- V(  5 ) )* V(  6)   R1=(1-L)*RM

1.000E-01 1.843E-04 3.112E-01 2.000E+00 1.032E+02
>R                    [ R for RUN to start fitting, FISH first asks for
                       integration scheme numbers, we enter -11 to test]
INTEGRATION SCHEME FOR I(Q) K6 =  (*)
( 0-7, -VE TO LIST WEIGHTS, -11 TO TEST ALL, SET K6<0 TO GET HERE )
-11
TEST AT Q= ?  (*)
.1
     Method     F**2(Q)        F(Q)          SUMX
  0  SIMPSON     2.284441E+13 -8.243047E+07  1.210804E+04
  1  GAUSS  4    1.850529E+13 -8.009058E+07  1.210425E+04
  2  GAUSS 10    2.285260E+13 -8.242504E+07  1.210481E+04
  3  RT+LN 10    2.285226E+13 -8.242519E+07  1.210481E+04
  4  GEN   10    2.279486E+13 -8.242452E+07  1.210480E+04
```

```
  5  GAUSS 16    2.285261E+13 -8.242503E+07  1.210481E+04
  6  GAUSS 32    2.285261E+13 -8.242500E+07  1.210481E+04
  7  GAUSS 64    2.285261E+13 -8.242500E+07  1.210481E+04
INTEGRATION SCHEME FOR I(Q) K6 =  (*)
( 0-7, -VE TO LIST WEIGHTS, -11 TO TEST ALL, SET K6<0 TO GET HERE )
7
INTEGRATION SCHEME FOR MOMENTS OF P(R), K7= (*)
-11
  0  SIMPSON   V=  6.8530E+09 SIG*2=  1.7402E+02 RBAR=  4.7706E+01
               ENT=  3.8283E+00
  1  GAUSS  4  V=  6.8504E+09 SIG*2=  1.7389E+02 RBAR=  4.7704E+01
               ENT=  1.0824E+00
  2  GAUSS 10  V=  6.8529E+09 SIG*2=  1.7402E+02 RBAR=  4.7706E+01
               ENT=  1.9536E+00
  3  RT+LN 10  V=  6.8529E+09 SIG*2=  1.7402E+02 RBAR=  4.7706E+01
               ENT=  1.7930E+00
  4  GEN   10  V=  6.8529E+09 SIG*2=  1.7402E+02 RBAR=  4.7706E+01
               ENT=  1.5818E+00
  5  GAUSS 16  V=  6.8529E+09 SIG*2=  1.7402E+02 RBAR=  4.7706E+01
               ENT=  2.4080E+00
  6  GAUSS 32  V=  6.8529E+09 SIG*2=  1.7402E+02 RBAR=  4.7706E+01
               ENT=  3.0869E+00
  7  GAUSS 64  V=  6.8529E+09 SIG*2=  1.7402E+02 RBAR=  4.7706E+01
               ENT=  3.7725E+00
INTEGRATION SCHEME FOR MOMENTS OF P(R), K7= (*)
7                    [now it continues with proper calc]
 POLYDISPERSE R=  3.00E+01 TO  9.00E+01  N=   1 TO  31
 DR=N(4)=  2.00E+00 RMAX=N(5)=  9.00E+01
 PMIN= 3.803E-14 PMAX= 1.250E-11 ENT= 3.773E+00 B=N(7)= 0.000E+00
 VNORM= 1.84280E-04  SIG(R)/RB= 2.7652E-01  SIG**2= 1.7402E+02
 RBAR= 4.7706E+01  AREA P(R)= 3.2551E-10
WEIGHT FUNCTION TYPE (K1=)   1,   IF K1=1, WTS = 1/SIGMA**2 , OTHERWISE
UNIT
WTS
CYC  1   49 DATA+ 0 PRED,  3 PAR  SWSE= 9.256E+03 XDWE= 9.171E+03 VAR=
1.928E+02
>1 6                 [ look at first 6 parameters]
 1  6 81 R**N VOL     1.842795E-04    1.789E-04   1.0 -8.65E-05
 2  6  2 R2-R1        6.000000E+01    0.000E+00  -1.0  0.00E+00
```

```
 3  6  3 R1           3.000000E+01    0.000E+00  -1.0  0.00E+00
 4  6  4 N           -1.500000E+00    0.000E+00   0.0  0.00E+00
 5  2  1 LAMBDA       5.000000E-01    0.000E+00   0.0  0.00E+00
 6  2  2    R-MID     6.000000E+01    1.238E+01   1.0  2.57E+01
>R                  [ run another cycle]
 POLYDISPERSE R=  4.28E+01 TO  1.28E+02  N=   1 TO  43
 DR=N(4)=  2.00E+00 RMAX=N(5)=  1.28E+02
 PMIN= 7.315E-15 PMAX= 1.656E-12 ENT= 3.765E+00 B=N(7)= 0.000E+00
 VNORM= 9.77584E-05  SIG(R)/RB= 2.7495E-01  SIG**2= 3.4823E+02
 RBAR= 6.7872E+01  AREA P(R)= 6.0102E-11
CYC  2   49 DATA+ 0 PRED,  3 PAR  SWSE= 1.068E+04 XDWE= 1.060E+04 VAR=
2.226E+02          [ fit even worse]
>P
 1  6 81 R**N VOL     9.775841E-05    2.330E-04   1.0  9.46E-04
 2  6  2 R2-R1        8.566107E+01    0.000E+00  -1.0  0.00E+00
 3  6  3 R1           4.283054E+01    0.000E+00  -1.0  0.00E+00
 4  6  4 N           -1.500000E+00    0.000E+00   0.0  0.00E+00
 5  2  1 LAMBDA       5.000000E-01    0.000E+00   0.0  0.00E+00
 6  2  2    R-MID     8.566107E+01    3.213E+01   1.0 -8.31E+01
 7  3  1 BKG A        1.364664E+00    6.109E-01   1.0  4.34E-02
 8  3  2     B*Q      0.000000E+00    0.000E+00   0.0  0.00E+00
 9  3  3     C*Q**2   0.000000E+00    0.000E+00   0.0  0.00E+00
10 99  1    SCALE     1.000000E+00    0.000E+00   0.0  0.00E+00
> 6=F=0.4           [ damp down shifts on V6 and V1 as they are moving too much ]
 6  2  2    R-MID     8.566107E+01    3.213E+01   0.4 -8.31E+01
> 1=N=0.4
 1  6 81 R**N VOL     9.775841E-05    2.330E-04   0.4  9.46E-04
>R
 POLYDISPERSE R=  2.62E+01 TO  7.86E+01  N=   1 TO  27
 DR=N(4)=  2.00E+00 RMAX=N(5)=  7.86E+01
 PMIN= 2.675E-13 PMAX= 5.451E-11 ENT= 3.776E+00 B=N(7)= 0.000E+00
 VNORM= 4.75972E-04  SIG(R)/RB= 2.7726E-01  SIG**2= 1.3389E+02
 RBAR= 4.1735E+01  AREA P(R)= 1.2544E-09
CYC  3   49 DATA+ 0 PRED,  3 PAR  SWSE= 3.971E+03 XDWE= 3.355E+03 VAR=
8.273E+01
>1 7
 1  6 81 R**N VOL     4.759722E-04    1.121E-04   0.4 -5.99E-04
 2  6  2 R2-R1        5.240216E+01    0.000E+00  -1.0  0.00E+00
 3  6  3 R1           2.620108E+01    0.000E+00  -1.0  0.00E+00
```

```
 4   6   4 N           -1.500000E+00    0.000E+00   0.0  0.00E+00
 5   2   1 LAMBDA       5.000000E-01    0.000E+00   0.0  0.00E+00
 6   2   2    R-MID     5.240216E+01    2.821E+00   0.4  1.28E+01
 7   3   1 BKG A        1.408045E+00    3.879E-01   1.0  3.27E-02
>R
 POLYDISPERSE R=  2.88E+01 TO  8.63E+01  N=   1 TO  29
 DR=N(4)=  2.00E+00 RMAX=N(5)=  8.63E+01
 PMIN= 1.426E-13 PMAX= 1.888E-11 ENT= 3.774E+00 B=N(7)= 0.000E+00
 VNORM= 2.36320E-04  SIG(R)/RB= 2.7674E-01  SIG**2= 1.6032E+02
 RBAR= 4.5753E+01  AREA P(R)= 4.7304E-10
CYC  4   49 DATA+ 0 PRED,  3 PAR  SWSE= 3.888E+03 XDWE= 3.723E+03 VAR=
8.099E+01
> PLOT                    [ still getting nowhere, do a plot to see fit]
 PLOT CONTROLS  IDEV= 0-EXIT, 1-screen, 2-file
error bars are IEB*sigma (ONLY USE ON LINEAR PLOT )
IPW=1 adds scaled wts
IDEL spreads graphs apart (IDEL=2 is default)
     IDEV,IEB,IPW,IDEL =        (5I1)
11
 OVER-PLOT P(Q)  ?  ( LTYP,LSYM (2I1) )          [return to enter zero]
 OVER-PLOT RESCALED S(Q)  ? ( LTYP,LSYM (2I1) )
 OVER-PLOT RESCALED BETA(Q)  ? ( LTYP,LSYM (2I1) )
SUBTRACT BACKGROUND ("WRK",MODELS 3&4) ? (ANS 1)

       AUTO      CHOICE
X1   0.000E+00 0.000E+00
X2   2.258E-01 2.258E-01
Y1  -7.411E+02-7.411E+02
Y2   8.725E+02 8.725E+02

AXES  1-AUTO, 2-U CHOOSE, 3-USE CHOICE, 0-RETURN ?  1
TITLE ?            [return to enter zero]
```

[you may be asked to check plot device type, plot appears on screen, lousy fit to data, try another model]

```
 PLOT CONTROLS  IDEV= 0-EXIT, 1-screen, 2-file
error bars are IEB*sigma (ONLY USE ON LINEAR PLOT )
IPW=1 adds scaled wts
IDEL spreads graphs apart (IDEL=2 is default)
```

```
     IDEV,IEB,IPW,IDEL =         (5I1)              [return to get out of plot]

TALK ROUTINE
>STOP
DID YOU REMEMBER TO PF AND FF ? ,NOW SAVE CALC SETS


1-READ MODEL FILE              4-CALC ONLY, CHOOSE Q
2-CHOOSE OBS,CALC SETS ETC.    5-INDEX
3-ENTER FIT ROUTINE            6-RETURN TO MAIN ROUTINE
1                   [ continue reading model file]
LSQIN READS MODEL DESCRIPTION FILE FROM FORTRAN STREAM 3
ENTER 0-TO IGNORE, 1-USE THIS MODEL, 9-REWIND
 (POLY DISP SPHERES)*HSS(Q)    CONSTRAINED TO FIT RM AND LAM
 1  6 81 R**N VOL     1.842795E-04    5.014E-06   1.0 -2.22E-04
 2  6  2 R2-R1        7.326303E+01    0.000E+00  -1.0  0.00E+00
 3  6  3 R1           2.997124E+01    0.000E+00  -1.0  0.00E+00
 4  6  4 N           -1.500000E+00    0.000E+00   0.0  0.00E+00
USE=1               [ use this one, same as before but with hard sphere structure factor added]
1-READ MODEL FILE              4-CALC ONLY, CHOOSE Q
2-CHOOSE OBS,CALC SETS ETC.    5-INDEX
3-ENTER FIT ROUTINE            6-RETURN TO MAIN ROUTINE
3
>P
 1  6 81 R**N VOL     1.842795E-04    5.014E-06   1.0 -2.22E-04
 2  6  2 R2-R1        7.326303E+01    0.000E+00  -1.0  0.00E+00
 3  6  3 R1           2.997124E+01    0.000E+00  -1.0  0.00E+00
 4  6  4 N           -1.500000E+00    0.000E+00   0.0  0.00E+00
 5  2  1 LAMBDA       5.500000E-01    0.000E+00   0.0  0.00E+00
 6  2  2    R-MID     6.660275E+01    4.136E-01   1.0  2.30E+01
 7 22  1 HS S(Q) VOL  1.854555E-01    1.444E-02   1.0 -6.66E-02
 8 22  2 SPH RADIUS   1.351837E+02    2.197E+00   0.4  0.00E+00
 9  3  1 BKG A        9.848869E-02    1.601E-02   1.0  0.00E+00
10  3  2     B*Q      0.000000E+00    0.000E+00   0.0  0.00E+00
11  3  3     C*Q**2   0.000000E+00    0.000E+00   0.0  0.00E+00
12 15  1 SMEAR        1.000000E+00    0.000E+00   0.0  0.00E+00
13 15  2 NSIMP        2.100000E+01    0.000E+00   0.0  0.00E+00
14 99  1    SCALE     1.000000E+00    0.000E+00   0.0  0.00E+00
> OFF               [ trick to zero the shifts column]
```

```
> ON
> R
 POLYDISPERSE R=  3.00E+01 TO  1.03E+02  N=   1 TO  37
 DR=N(4)=  2.00E+00 RMAX=N(5)=  1.03E+02
 PMIN= 3.686E-14 PMAX= 7.843E-12 ENT= 3.768E+00 B=N(7)= 0.000E+00
 VNORM= 1.84280E-04  SIG(R)/RB= 3.1124E-01  SIG**2= 2.5661E+02
 RBAR= 5.1469E+01  AREA P(R)= 2.4582E-10
SMEARING ROUTINE CALLED

NPSMEAR,NSHAPE,SCALE,NSIMP =   12    0     1.000   21
NOTE for debug purposes, non-smeared calc is in SET 5
CYC  5   49 DATA+ 0 PRED,  5 PAR  SWSE= 9.930E+03 XDWE= 9.916E+03 VAR=
2.159E+02
>1 9
 1  6 81 R**N VOL     1.842795E-04    2.491E-04   1.0  1.32E-04
 2  6  2 R2-R1        7.326303E+01    0.000E+00  -1.0  0.00E+00
 3  6  3 R1          2.997124E+01    0.000E+00  -1.0  0.00E+00
 4  6  4 N          -1.500000E+00    0.000E+00   0.0  0.00E+00
 5  2  1 LAMBDA       5.500000E-01    0.000E+00   0.0  0.00E+00
 6  2  2    R-MID     6.660275E+01    1.813E+01   1.0  1.40E+01
 7 22  1 HS S(Q) VOL  1.854555E-01    7.833E-01   1.0 -2.72E-01
 8 22  2 SPH RADIUS   1.351837E+02    1.690E+02   0.4 -4.32E+01
 9  3  1 BKG A        9.848869E-02    6.262E-01   1.0  1.21E+00
> R
 POLYDISPERSE R=  3.63E+01 TO  1.25E+02  N=   1 TO  45
 DR=N(4)=  2.00E+00 RMAX=N(5)=  1.25E+02
 PMIN= 1.665E-14 PMAX= 6.382E-12 ENT= 3.764E+00 B=N(7)= 0.000E+00
 VNORM= 3.15856E-04  SIG(R)/RB= 3.1045E-01  SIG**2= 3.7255E+02
 RBAR= 6.2173E+01  AREA P(R)= 2.3933E-10
SMEARING ROUTINE CALLED

NPSMEAR,NSHAPE,SCALE,NSIMP =   12    0     1.000   21
NOTE for debug purposes, non-smeared calc is in SET 5
CYC  6   49 DATA+ 0 PRED,  5 PAR  SWSE= 3.126E+03 XDWE= 3.080E+03 VAR=
6.795E+01
>1 9
 1  6 81 R**N VOL     3.158562E-04    1.902E-04   1.0  5.24E-04
 2  6  2 R2-R1        8.866681E+01    0.000E+00  -1.0  0.00E+00
 3  6  3 R1          3.627278E+01    0.000E+00  -1.0  0.00E+00
```

```
 4  6  4 N           -1.500000E+00    0.000E+00   0.0  0.00E+00
 5  2  1 LAMBDA       5.500000E-01    0.000E+00   0.0  0.00E+00
 6  2  2    R-MID     8.060619E+01    9.750E+00   1.0 -1.46E+01
 7 22  1 HS S(Q) VOL -8.685419E-02    1.518E-01   1.0  1.32E-01
 8 22  2 SPH RADIUS   1.179190E+02    1.564E+02   0.4  5.50E+01
 9  3  1 BKG A        1.306350E+00    3.447E-01   1.0  2.15E-02
>R
 POLYDISPERSE R=  2.97E+01 TO  1.02E+02  N=   1 TO  37
 DR=N(4)=  2.00E+00 RMAX=N(5)=  1.02E+02
 PMIN= 1.232E-13 PMAX= 3.711E-11 ENT= 3.768E+00 B=N(7)= 0.000E+00
 VNORM= 8.40031E-04  SIG(R)/RB= 3.1128E-01  SIG**2= 2.5185E+02
 RBAR= 5.0983E+01  AREA P(R)= 1.1528E-09
SMEARING ROUTINE CALLED

NPSMEAR,NSHAPE,SCALE,NSIMP =   12    0     1.000   21
NOTE for debug purposes, non-smeared calc is in SET 5
CYC  7   49 DATA+ 0 PRED,  5 PAR  SWSE= 8.133E+03 XDWE= 8.118E+03 VAR=
1.768E+02
>1 9                [ the fit OUGHT to converge as XDWE is same as SWSE]
 1  6 81 R**N VOL     8.400314E-04    2.212E-04   1.0 -5.39E-04
 2  6  2 R2-R1        7.256344E+01    0.000E+00  -1.0  0.00E+00
 3  6  3 R1          2.968504E+01    0.000E+00  -1.0  0.00E+00
 4  6  4 N           -1.500000E+00    0.000E+00   0.0  0.00E+00
 5  2  1 LAMBDA       5.500000E-01    0.000E+00   0.0  0.00E+00
 6  2  2    R-MID     6.596676E+01    3.524E+00   1.0  3.24E+00
 7 22  1 HS S(Q) VOL  4.478095E-02    1.780E-01   1.0  2.13E-02
 8 22  2 SPH RADIUS   1.399234E+02    2.009E+02   0.4 -5.90E+01
 9  3  1 BKG A        1.327847E+00    5.664E-01   1.0 -2.61E-02
> 1=N=.5            [ damp down V1 which is oscillating]
 1  6 81 R**N VOL     8.400314E-04    2.212E-04   0.5 -5.39E-04
>R
 POLYDISPERSE R=  3.11E+01 TO  1.07E+02  N=   1 TO  39
 DR=N(4)=  2.00E+00 RMAX=N(5)=  1.07E+02
 PMIN= 4.870E-14 PMAX= 2.092E-11 ENT= 3.767E+00 B=N(7)= 0.000E+00
 VNORM= 5.70646E-04  SIG(R)/RB= 3.1107E-01  SIG**2= 2.7650E+02
 RBAR= 5.3456E+01  AREA P(R)= 6.7962E-10
SMEARING ROUTINE CALLED

NPSMEAR,NSHAPE,SCALE,NSIMP =   12    0     1.000   21
```

```
NOTE for debug purposes, non-smeared calc is in SET 5
CYC  8   49 DATA+ 0 PRED,  5 PAR  SWSE= 2.685E+02 XDWE= 2.487E+02 VAR=
5.836E+00
> 1 9
 1  6 81 R**N VOL     5.706458E-04    4.494E-05   0.5 -1.47E-04
 2  6  2 R2-R1        7.612236E+01    0.000E+00  -1.0  0.00E+00
 3  6  3 R1           3.114096E+01    0.000E+00  -1.0  0.00E+00
 4  6  4 N           -1.500000E+00    0.000E+00   0.0  0.00E+00
 5  2  1 LAMBDA       5.500000E-01    0.000E+00   0.0  0.00E+00
 6  2  2    R-MID     6.920214E+01    1.101E+00   1.0  2.25E+00
 7 22  1 HS S(Q) VOL  6.605905E-02    3.671E-02   1.0  4.61E-02
 8 22  2 SPH RADIUS   1.163301E+02    2.390E+01   0.4  3.97E+00
 9  3  1 BKG A        1.301789E+00    1.027E-01   1.0 -2.45E-03
> R
 POLYDISPERSE R=  3.22E+01 TO  1.11E+02  N=   1 TO  40
 DR=N(4)=  2.00E+00 RMAX=N(5)=  1.11E+02
 PMIN= 4.823E-14 PMAX= 1.610E-11 ENT= 3.766E+00 B=N(7)= 0.000E+00
 VNORM= 4.97388E-04  SIG(R)/RB= 3.1093E-01  SIG**2= 2.9429E+02
 RBAR= 5.5172E+01  AREA P(R)= 5.3890E-10
SMEARING ROUTINE CALLED

NPSMEAR,NSHAPE,SCALE,NSIMP =   12    0     1.000   21
NOTE for debug purposes, non-smeared calc is in SET 5
CYC  9   49 DATA+ 0 PRED,  5 PAR  SWSE= 6.053E+01 XDWE= 4.033E+01 VAR=
1.316E+00
> 1
 1  6 81 R**N VOL     4.973885E-04    2.178E-05   0.5  1.52E-05
>R
 POLYDISPERSE R=  3.23E+01 TO  1.11E+02  N=   1 TO  40
 DR=N(4)=  2.00E+00 RMAX=N(5)=  1.11E+02
 PMIN= 5.800E-14 PMAX= 1.606E-11 ENT= 3.766E+00 B=N(7)= 0.000E+00
 VNORM= 5.05004E-04  SIG(R)/RB= 3.1091E-01  SIG**2= 2.9691E+02
 RBAR= 5.5421E+01  AREA P(R)= 5.3984E-10
SMEARING ROUTINE CALLED

NPSMEAR,NSHAPE,SCALE,NSIMP =   12    0     1.000   21
NOTE for debug purposes, non-smeared calc is in SET 5
CYC 10   49 DATA+ 0 PRED,  5 PAR  SWSE= 4.341E+01 XDWE= 2.320E+01 VAR=
9.438E-01
```

```
> 1=N
 1  6 81 R**N VOL     5.050041E-04    1.850E-05   1.0  2.03E-05
> R
 POLYDISPERSE R=  3.23E+01 TO  1.11E+02  N=   1 TO  40
 DR=N(4)=  2.00E+00 RMAX=N(5)=  1.11E+02
 PMIN= 6.004E-14 PMAX= 1.672E-11 ENT= 3.766E+00 B=N(7)= 0.000E+00
 VNORM= 5.25346E-04  SIG(R)/RB= 3.1091E-01  SIG**2= 2.9683E+02
 RBAR= 5.5413E+01  AREA P(R)= 5.6182E-10
SMEARING ROUTINE CALLED

NPSMEAR,NSHAPE,SCALE,NSIMP =   12    0     1.000   21
NOTE for debug purposes, non-smeared calc is in SET 5
CYC 11   49 DATA+ 0 PRED,  5 PAR  SWSE= 2.022E+01 XDWE= 8.952E-03 VAR=
4.396E-01
> R
 POLYDISPERSE R=  3.23E+01 TO  1.11E+02  N=   1 TO  40
 DR=N(4)=  2.00E+00 RMAX=N(5)=  1.11E+02
 PMIN= 5.997E-14 PMAX= 1.671E-11 ENT= 3.766E+00 B=N(7)= 0.000E+00
 VNORM= 5.24995E-04  SIG(R)/RB= 3.1091E-01  SIG**2= 2.9682E+02
 RBAR= 5.5412E+01  AREA P(R)= 5.6147E-10
SMEARING ROUTINE CALLED

NPSMEAR,NSHAPE,SCALE,NSIMP =   12    0     1.000   21
NOTE for debug purposes, non-smeared calc is in SET 5
CYC 12   49 DATA+ 0 PRED,  5 PAR  SWSE= 2.021E+01 XDWE= 3.490E-04 VAR=
4.394E-01
> R
 POLYDISPERSE R=  3.23E+01 TO  1.11E+02  N=   1 TO  40
 DR=N(4)=  2.00E+00 RMAX=N(5)=  1.11E+02
 PMIN= 5.998E-14 PMAX= 1.670E-11 ENT= 3.766E+00 B=N(7)= 0.000E+00
 VNORM= 5.24919E-04  SIG(R)/RB= 3.1091E-01  SIG**2= 2.9682E+02
 RBAR= 5.5413E+01  AREA P(R)= 5.6137E-10
SMEARING ROUTINE CALLED

NPSMEAR,NSHAPE,SCALE,NSIMP =   12    0     1.000   21
NOTE for debug purposes, non-smeared calc is in SET 5
CYC 13   49 DATA+ 0 PRED,  5 PAR  SWSE= 2.021E+01 XDWE= 3.399E-04 VAR=
4.394E-01
> PP                [ fit converged, XDWE is small compared to SWSE ]
```

```
T  1 P 14 S  1 C  2 N  5
W  1 K  0 IP 1 MS 1 IY 1 Q  7 R  7    6    1    1    0    2    0   49
13    5
 (POLY DISP SPHERES)*HSS(Q)    CONSTRAINED TO FIT RM AND LAM
 1  6 81 R**N VOL     5.249193E-04    1.261E-05   1.0  7.24E-08
 2  6  2 R2-R1        7.893859E+01    0.000E+00  -1.0  0.00E+00
 3  6  3 R1           3.229306E+01    0.000E+00  -1.0  0.00E+00
 4  6  4 N           -1.500000E+00    0.000E+00   0.0  0.00E+00
 5  2  1 LAMBDA       5.500000E-01    0.000E+00   0.0  0.00E+00
 6  2  2    R-MID     7.176235E+01    3.436E-01   1.0 -1.19E-03
 7 22  1 HS S(Q) VOL  1.150149E-01    1.060E-02   1.0  4.80E-05
 8 22  2 SPH RADIUS   1.183947E+02    3.820E+00   0.4  8.52E-02
 9  3  1 BKG A        1.297392E+00    2.811E-02   1.0 -4.03E-06
10  3  2     B*Q      0.000000E+00    0.000E+00   0.0  0.00E+00
11  3  3     C*Q**2   0.000000E+00    0.000E+00   0.0  0.00E+00
12 15  1 SMEAR        1.000000E+00    0.000E+00   0.0  0.00E+00
13 15  2 NSIMP        2.100000E+01    0.000E+00   0.0  0.00E+00
14 99  1    SCALE     1.000000E+00    0.000E+00   0.0  0.00E+00
  1  1 2658         CALC 2 BKG 0 POL 3  SSE= 2.021E+01
CONSTRAINT  1 TYPE  3 USE= 1
V(  2 ) = 2.0* V(  5 ) * V(  6 )   R2-R1=2*L *RM

CONSTRAINT  2 TYPE  4 USE= 1
V(  3 ) = (1.0- V(  5 ) )* V(  6)   R1=(1-L)*RM

1.000E-01 5.249E-04 3.109E-01 2.000E+00 1.112E+02
>PLOT
 PLOT CONTROLS  IDEV= 0-EXIT, 1-screen, 2-file
error bars are IEB*sigma (ONLY USE ON LINEAR PLOT )
IPW=1 adds scaled wts
IDEL spreads graphs apart (IDEL=2 is default)
     IDEV,IEB,IPW,IDEL =        (5I1) 11         [ to screen with errors]
 OVER-PLOT P(Q)  ?  ( LTYP,LSYM (2I1) ) 1        [ solid line]
 OVER-PLOT RESCALED S(Q)  ? ( LTYP,LSYM (2I1) ) 54          [ xxx]
 OVER-PLOT RESCALED BETA(Q)  ? ( LTYP,LSYM (2I1) ) 3        [ dashed]
SUBTRACT BACKGROUND ("WRK",MODELS 3&4) ? (ANS 1) 0
       AUTO      CHOICE
X1   0.000E+00 0.000E+00
X2   2.258E-01 2.258E-01
```

```
Y1  -2.900E+02-7.411E+02
Y2   8.620E+02 8.725E+02

AXES  1-AUTO, 2-U CHOOSE, 3-USE CHOICE, 0-RETURN ? 1
TITLE ?
 EXAMPLE FIT TO RUN 2658 WITH POLYDISP INTERACTING SPHERES
```

[plot 1 appears here]

```
 S(Q) *  5.00E+02 +   0.000E+00
```
[rescale and shift used to overplot S(Q) on same axes]

```
 BETA(Q) *  5.00E+02 +   0.000E+00

 PLOT CONTROLS  IDEV= 0-EXIT, 1-screen, 2-file
error bars are IEB*sigma (ONLY USE ON LINEAR PLOT )
IPW=1 adds scaled wts
IDEL spreads graphs apart (IDEL=2 is default)
     IDEV,IEB,IPW,IDEL =        (5I1)
```
[return to leave plot]

```
TALK ROUTINE
> STOP
```
[go back to FIT menu]

```
DID YOU REMEMBER TO PF AND FF ? ,NOW SAVE CALC SETS


1-READ MODEL FILE              4-CALC ONLY, CHOOSE Q
2-CHOOSE OBS,CALC SETS ETC.    5-INDEX
3-ENTER FIT ROUTINE            6-RETURN TO MAIN ROUTINE
6
```
[go back to main program]

```
Command> LIST
 WRITE SET I TO DATA FILE (J=1), MONITOR(J=2) OR TO
SCREEN(J=0)  I,J=? (2I1) 21
```
[save calculated I(Q) in a file]

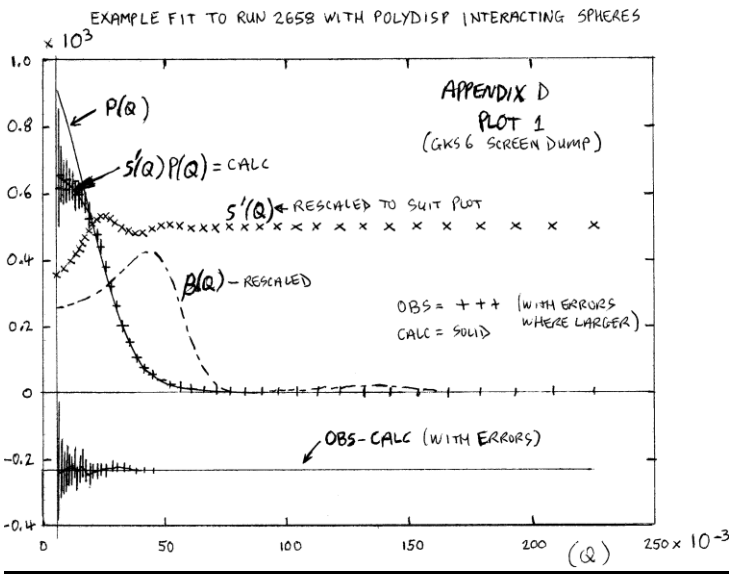
```
SET( 2 )   CAL
OLD TITLES:
                         CAL
                         CAL
TITLE:
26-MAY-89     09:24:22   CAL
```

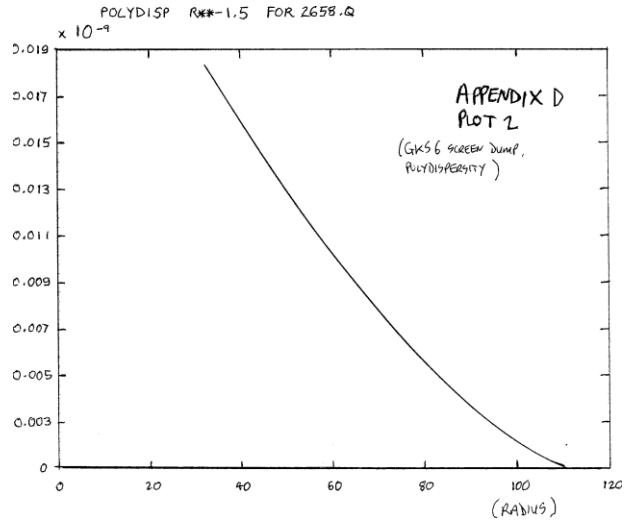
```
                        CAL
 0-CONTINUE, 1-CHANGE TITLES, 2-RETURN  ? 1
ENTER TWO TITLE RECORDS (14A4, 20A4) :
    FIT TO 2658.Q
    second title line

TITLE:
26-MAY-89     09:24:22      FIT TO 2658.Q

 0-CONTINUE, 1-CHANGE TITLES, 2-RETURN  ? 0
Command>PLOT              [ use the long plot routine to draw the P(R) in set 3]
Welcome to the PLOT routine, enter HELP after
          Plot>    if you are lost
Plot>READ
HOW MANY CURVES ? (I1) 1
FOR EACH CURVE TO BE DRAWN ENTER ON THREE LINES :
SET NUMBER (I1)
LTYPE,LSYMBOL,IFOLD,IEB,INUM-NOT-Q   (5I1)
Y SHIFT (F12)
CURVE 1 SET ? (I1) 3
LT LSY IFLD IEB INNQ   (5I1) 1
YSHIFT 0.0
Plot> SCREEN
       AUTO      CHOICE
X1   0.000E+00 0.000E+00
X2   1.103E+02 2.258E-01
Y1   1.200E-13-7.411E+02
Y2   1.734E-11 8.725E+02

AXES  1-AUTO, 2-U CHOOSE, 3-USE CHOICE, 0-RETURN ? 1
TITLE ?
 POLYDISP  R**-1.5  FOR 2658.Q
```

[plot 2 appears here]

```
Plot> STOP
Command> STOP
FORTRAN STOP
$
```

$ [returns to VMS operating system]





Appendix E - Least Squares and Marquardt Fits

General non-linear least squares fitting, of the sort required for most SANS data, relies on some understanding of statistical distributions, estimation of functions and solving a set of equations (matrix inversion). Thus the subject usually appears towards the end of text books on "numerical methods" ! All that I can do here is try to portray very briefly the route that is followed. Chapter 15 of "Numerical Recipes" gives a very much more rigorous account. ["Numerical Recipes in FORTRAN, The art of scientific computing", W.H.Press, S.A.Teukolsky, W.T.Vetterling & B.P.Flannery, Cambridge University Press, 2nd Edition, 1992, reprinted 1994. Other versions of the book & software CD are available for Pascal, C, & Basic.]

Suppose our data are y_i where i = 1 to N at points x_i - these could be SANS intensities at N values of scattering vector Q value.

Let the parameters in our model be a_j where j = 1 to M

Calculated data points are $CALC_i$ = function(x_i, a_1, a_2, … a_M)

Define a "merit function" $$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - CALC_i}{\sigma_i} \right)^2 \qquad (1)$$

If the errors (standard deviations) on data σ_i are *independent* and have a *normal (Gaussian) distribution* then statistical theories tell us that the minimum in χ^2 is the "most likely" solution and to expect that $\chi^2/(N-M) \sim 1$.

(A normal distribution is within $\pm 2\sigma$ 68% of the time, $\pm 3\sigma$ 95% of the time. The "Poisson distribution" for neutron counts has a broader tail for small counts. Since the merit function is then not quite correct, "outliers" can be a problem, set their W_i to zero ??)

At the minimum of χ^2 its derivative with respect to each of the parameters a_j will be zero:

$$\frac{\partial(\chi^2)}{\partial a_j} = 0 \quad \text{gives } M \text{ equations } \quad \sum_i W_i (y_i - CALC_i) \frac{\partial CALC_i}{\partial a_j} = 0 \qquad (2)$$

Where weights $W_i = 1/\sigma_i^2$

For the simplest case the model is "linear" in the parameters, each of which is just a scale factor in front of some mathematical function (which itself may be very non-linear) so that:

$$CALC_i = \sum_{k=1}^{M} a_k Fun_k(x_i) = \sum_{k=1}^{M} a_k D_{ik} \qquad (3)$$

Note though that each "basis function" is actually just the derivative D_{ij} of the calculated model for that parameter:

$$D_{ij} = \frac{\partial CALC_i}{\partial a_j} = Fun_j(x_i) \qquad (4)$$

In the linear case the set of equations (2) can be solved exactly to give the parameter values a_j. We will now however make the equations slightly more complicated by anticipating the method for the more general "non-linear" case !

Assume the present parameters a_j^{now} give $CALC_i^{now}$ with differences $E_i = (y_i - CALC_i^{now})$

We need to shift the parameters to $a_j = a_j^{now} + \Delta a_j$ to give the best (or at least a smaller) value of χ^2. Since the problem is linear, we can use (3) to write

$$CALC_i = CALC_i^{now} + \sum_k \Delta a_k D_{ik} \qquad (5)$$

which is substituted into (2) to give a set of M equations:

$$\sum_i W_i \left(E_i - \sum_k \Delta a_k D_{ik} \right) D_{ij} = 0 \qquad (6)$$

The equations are easier to manipulate in matrix form as:

$$\Delta a (D^T W D) - D^T W E = 0$$

which may be rearranged to give the desired a column of parameter shifts Δa as a product of a square "least squares matrix" and a column matrix :

$$\Delta a = (D^T W D)^{-1}(D^T W E) \qquad (7)$$

Δa is a column vector of M rows, derivative matrix D has N rows x M columns, weights W_{ii} are a diagonal N x N, and differences E is a column of N rows. Superscripts "T" mean transpose (i.e. $D_{ij}^T = D_{ji}$) and "-1" means the matrix inverse.

[If you are not familiar with this kind of notation just concentrate on the meaning of the results ! If an example might also help, then for the case of just two parameter and 3 data points, the simultaneous equations to be solved to find shifts $a = \Delta a_1$ & $b = \Delta a_2$ are:

$$a(W_1 D_{1a}{}^2 + W_2 D_{2a}{}^2 + W_3 D_{3a}{}^2) + b(W_1 D_{1a} D_{1b} + W_2 D_{2a} D_{2b} + W_3 D_{3a} D_{3b}) - (W_1 D_{1a} E_1 + W_2 D_{2a} E_2 + W_3 D_{3a} E_3) = 0$$
$$a(W_1 D_{1a} D_{1b} + W_2 D_{2a} D_{2b} + W_3 D_{3a} D_{3b}) + b(W_1 D_{1b}{}^2 + W_2 D_{2b}{}^2 + W_3 D_{3b}{}^2) - (W_1 D_{1b} E_1 + W_2 D_{2b} E_2 + W_3 D_{3b} E_3) = 0$$

]

In a "linear case" (such as a polynomial or straight line fit) Δa gives an immediate and exact solution, even with zero starting parameters.

For a "non-linear" case (as almost all SANS) equations (3) to (5) are only *approximately* true, so the solution must be iterated. Note that the derivatives D_{ij} are now no longer constant at each Q value, so they must be recalculated for each iteration.

Fortunately it can be shown that ignoring the second derivatives in (3) is not detrimental, since they are usually small and statistically they should tend to cancel out when summed over the data. Nor does this have an effect on the location of the χ^2 minimum, only on the route taken to reach it.

It is a remarkable fact that the least squares method works at all for the "non-linear" cases found in SANS !

M x M matrix $C = (D^T W D)^{-1}$ is the variance-covariance matrix, its diagonal elements are $C_{jj} = \sigma_j^2$ the square of the *statisitical* standard deviation for each parameter a_j - *assuming* the conditions mentioned above on σ_i for the data are valid ! Effects of systematic errors, such as imperfect data treatment, or say Q resolution not explicitly included in the model must be considered separately !

Frequently the parameter "errors" obtained for SANS data are unrealistically small, as some trial and error adjustments and common sense will easily show. Off-diagonal elements C_{ij} give the correlation coefficients between parameters, which are helpful to identify poor parametrisation of a model.

Note there is no standard notation for the different matrices (e.g. derivatives D are sometimes called J for Jacobian). In some applications it may be useful to include off-diagonal elements in the weight matrix W to allow for "correlation" between data points. This can help to give more realistic error estimates.

Practical results

All that is needed for least squares fits are (a) routines to calculate the model CALC and its derivatives D for a given set of parameters and Q values and (b) a routine to invert a symmetric matrix (i.e. to solve a set of equations).

Derivative D_{ij} may be calculated numerically (but less reliably) by simply shifting the value of a parameter a_j temporarily by a small amount δ and calling the model routine again:

$$D_{ij} \approx \frac{1}{\delta}\Big(CALC(x_i, a_1, a_2, ..(a_j + \delta), ...a_n) - CALC(x_i, a_1, a_2, ..a_j, ...a_n)\Big)$$

The "non-linear" nature of most SANS problems requires that the least squares solutions are iterated. In a well behaved system each iteration gets closer to the χ^2 minimum. When χ^2 ceases to improve further the fit has "converged". Alas not all systems are "well behaved", so we must learn some ways to cope !

Steepest Descent & the Marquardt method

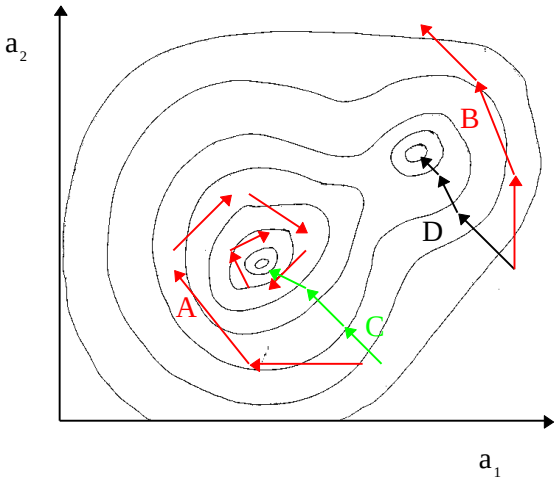
If the χ^2 merit function (1) is not well approximated by a quadratic near its minimum then the iterative least squares solution may not work.

Some oscillatory behaviour of given Δa_j between iterations may be *damped* down by applying only a fraction of the calculated shifts. Worse behaviour might require some time consuming trial & error on key parameters.

Following the "steepest descent" gradient of the χ^2 surface may be an alternative route to the desired minimum. Least squares tends to spiral down at right angles to the steepest descent of the χ^2 surface, thereby exploring more parameter space and being less likely to become stuck in a local minimum, but often in badly behaved cases "blowing up" altogether.

Marquardt (using an idea of Levenberg) noted a simple connection between the least squares and steepest descent routes. Replacing (D^TWD) by a constant diagonal gives the steepest descent route. By multiplying the diagonal elements of (D^TWD) by ($1+\lambda$) where λ is small for least squares or large for steepest descent gives a route that varies between the two extremes.

Schematic Least Square & Steepest Descent, on a 2 parameter χ^2surface



A - Well behaved least squares, explores reasonable parameter space.

B - Least squares "blows up" as shifts are too large (could be damped down ?)

C - Steepest descent from new starting point finds best fit.

D - Steepest descent finds a local minimum.

The Marquardt method would steer between B & D, but still might fall into the local minimum !

The Marquardt recipe:

(i) start with a modest $\lambda \sim 1$,

(ii) compute D (and save it) and χ^2

(iii) calculate parameter shifts using (7) with diagonal elements of (D^TWD) multiplied by ($1+\lambda$)

(iv) compute new parameters and their χ^2

(v) if fit has converged, or too many iterations, stop !

(vi) if fit improves, keep new parameters, divide λ by 10 and return to (ii)

(vii) if fit worsens, multiply λ by 10, return to (iii) (no new computation of D)

NOTE - to obtain the proper error estimates σ_j on parameters a_j set $\lambda=0$ for a final calculation.

The fit is guaranteed to improve, if only slowly, but not (in poorly behaved cases) to find a global minimum for χ^2 as the steepest descent route can become stuck in a "local minimum". It may be important, as with ordinary least squares, to try to find the solution again from slightly different starting points.

CONSTRAINTS

In many cases physical constraints and prior knowledge (e.g. shell to core molar volume ratio, consideration of fully extended surfactant tail lengths) may be required to locate *physically meaningful* parameters from amongst whole families of possible numerical solutions.

Absolute intensities are also vital either by constraining scale parameters to known volume fractions or concentrations or, where samples or intensities are less well known, by checking that fitted scale parameters are *consistent* with sample compositions.

If one parameter is constrained to another, then one could rewrite the model with one less parameter. This is not of course very convenient, so it is better to write the model with the maximum likely number of parameters that could be adjusted or investigated.

If parameter a_j is a function $f(a_k)$ of parameter a_k the model routine will give $\frac{\partial CALC_i}{\partial a_j}$ so the least squares calculation has to add an extra term to $\frac{\partial CALC_i}{\partial a_k}$ by:

$$\frac{\partial CALC_i}{\partial a_k} = \frac{\partial CALC_i}{\partial a_k} + \frac{\partial CALC_i}{\partial a_j}\frac{\partial a_j}{\partial a_k} = \frac{\partial CALC_i}{\partial a_k} + \frac{\partial CALC_i}{\partial a_j}\frac{\partial f(a_k)}{\partial a_k}$$

The FISH program will do this for a variety of pre-programmed constraint functions.

Appendix F - NUMERICAL INTEGRATION - GAUSSIAN QUADRATURES

Calculation of SANS intensities (and their derivatives) frequently requires some numerical integration, for example to sum over a polydisperse particle size, or in the form factors for monodisperse rods (discs) or ellipsoids.

Most of us will be familiar with the "trapezium rule" for integrating N data points, equally spaced *h* apart. Apart from the weights of ½ on the first and last points, this is just "adding up the data".

$$\int_{x_1}^{x_N} y(x)dx \approx h\left[\tfrac{1}{2}y_1 + y_2 + y_3 + \ldots + y_{N-1} + \tfrac{1}{2}y_N\right] \qquad (1)$$

Simpson's rule gives a better answer with the same points but different weights as it is equivalent to fitting a cubic equation through adjacent groups of three points.

$$\int_{x_1}^{x_N} y(x)dx \approx \frac{h}{3}\left[y_1 + 4y_2 + 2y_3 + 4y_4 \ \ldots \ + 2y_{N-2} + 4y_{N-1} + y_N\right] \qquad (2)$$

Removing the restriction on the points being *equally* spaced allows use of much more efficient "quadrature methods", which use a table of specially chosen x values (abscissae) and weights for the y values at those points. Since we are integrating the *model* function, the non-equally spaced points pose no problem! By some clever mathematics we may for example use a "10 point Gaussian quadrature" to integrate a function as if it were fitted by a 10^{th} order polynomial, or say a 64 point quadrature for an order 64 polynomial. The coefficients for a 10 point "Gauss-Legendre" quadrature are illustrated below. Note that though for symmetry the integration interval is here -1 to +1 the abscissae and weights can easily be rescaled to suit a given range.

$$\begin{aligned}\int_{-1}^{+1} y(x)dx \approx\ & 0.06667\big(y(-0.9739) + y(+0.9739)\big) + 0.14945\big(y(-0.8651) + y(+0.8651)\big) + \\ & + 0.21909\big(y(-0.6794) + y(+0.6794)\big) + 0.26923\big(y(-0.4334) + y(+0.4334)\big) + \\ & + 0.29552\big(y(-0.1489) + y(+0.1489)\big) \qquad (3)\end{aligned}$$

IF the function y(x) is *well* approximated by a polynomial then the resulting integral will be more accurate than say using Simpson's rule with many more points, and hence the fitting program will run more quickly. The abscissae in a Gaussian quadrature are grouped more closely towards the ends of the integration range.

(Actual integrations need the abscissae and weights to many more decimal places than illustrated in (3) above. Originally one looked them up in tables, but nowadays simple iterative routines are available to calculate them as needed. The precision of the computer becomes important for higher orders of quadrature.)

With modern computers the choice of integration scheme only makes a significant difference in speed for multi-dimensional integrals, such as for "oriented rods". In this latter case one must however be careful as the form factor for long, thin rods has strong oscillations for which Gaussian quadrature does not work well and Simpson's Rule is more appropriate for parts of the integration. It is advisable in any circumstance to test different integration schemes quite carefully, with increasing numbers of points to ensure that they converge adequately for a particular problem.

A full discussion of numerical integration methods (e.g. Chapter 4 in Numerical Recipes) focuses on issues such as estimation of accuracy and effects of singularities in the function. Certain functional singularities can be included in the quadrature, so that one "fits" say a polynomial times an exponential. These give rise to whole families of quadratures in addition to the usual "polynomial" Gauss-Legendre, such as Gauss-Chebyshev, Gauss-Laguerre, Gauss-Hermite and Gauss-Jacobi.

One advantage of the simple "trapezium rule" equation (1) is that is may efficiently be extended to include extra points between those of the previous set of points, to give an iterative answer with easily assessed "accuracy". (Various library routines are available in this area.)

The Gaussian quadrature method does not have this property, increasing the number of points results in a completely new set of abscissae and weights. The Gauss-Kronrod method is an "adaptive" integration scheme which expands the Gauss-Legendre polynomials in an optimal way, re-using results of the previous iteration. It uses a sequence of N points such as N=10, 21, 43, 87. Such "adaptive" schemes usually require to be given both an absolute and relative error expected for the result of the integration in order to know when to stop !