Fitting Hall Probe Measurements

Diane Grijio dos Santos(*)

CERN, Geneva, Switzerland

1. Mathematical Justification

1.1 Introduction

The project is concerned with fitting calibration curves to data obtained from Hall probes of magnetic fields.

We first make an exact fit using the "Uniform Cubic Spline" [1] and then use this to make a least squares Chebyshev fit.

1.2 Choice of cubic spline model

At the outset, the "rounded cubic spline model" [1] was tried on our data. This is a modification of the regular uniform cubic spline but is continuous in all derivatives.

The regular uniform cubic spline model can be written as

\[ y(x) = a_1 + a_2 x + a_3 x^2 + a_4 f(x-x_1) + \sum_{i=5}^{n} a_i f(x-x_{i-2}) \]

where \( f(x) = 0 \) if \( x < 0 \),

\( f(x) = x^3 \) if \( x > 0 \)

and \( n \) is the number of reference points.

For the rounded uniform cubic spline one uses

\[ f(x) = \left[ \frac{1}{a} \ln(1 + e^{\alpha x}) \right]^3, \]


DGS/ef
where α is a parameter to be chosen. For small α the function is very smooth, but the equations very ill-conditioned. For big α, the model is the same as the regular uniform cubic spline. However, with our data, even for α as large as 16, the equations were still ill-conditioned. In fact, when we used the spline coefficients, ai obtained this way, we found a noticeable ripple in the results which we suspected to be due to this ill-conditioning. With α as large as this, the function is, practically, identical to the regular cubic spline (fig. 1).

Indeed, when we subsequently used the regular uniform cubic spline, this ripple disappeared, the fit remaining otherwise the same (as we will see, this latter fit is less susceptible to rounding errors). We therefore, used the regular uniform cubic spline model.

1.3 The regular uniform cubic spline model

Let y(x) be the cubic spline model of the n reference points (xi, yi). The Taylor expansion of the model in the interval (xi, xi+1) then is

\[ y(x) = y_i + (x-x_i) y'_i + (x-x_i)^2 y''_i / 2 + (x-x_i)^3 (y'''_{i+1} - y'''_i) / (5D_i) \]

where \( D_i = x_{i+1} - x_i \) and \( y'_i \) is the derivative at \( x = x_i \).

For y(x) to be continuous, we have

\[ y_{i+1} = y_i + D_i y'_i + D_i^2 y''_i / 2 + D_i^3 (y'''_{i+1} - y'''_i) / 6 \]  \hspace{1cm} (2)

or

\[ y'_i = (y_{i+1} - y_i) / D_i - D_i y''_i / 3 - D_i^2 y'''_{i+1} / 6 . \]  \hspace{1cm} (3)

We consider \( y''_i \) as the unknowns and thus have \( n-2 \) equations with \( n \) unknowns

\[ D_i y''_i + 2 (D_i + D_{i+1}) y''_{i+1} + D_i y''_{i+2} = 6 \left( \frac{Y_{i+2} - Y_{i+1}}{D_i} - \frac{Y_{i+1} - Y_i}{D_i} \right), \]

where \( i = 1, n-2 \).

The two extra equations needed follow from the definition of the uniform cubic spline which is that \( y'''(x) \) is continuous at \( x_2 \) and \( x_{n-1} \) i.e.
\[
\frac{y''_3 - y''_2}{D_2} = \frac{y''_2 - y''_1}{D_1}
\]
and
\[
\frac{y''_n - y''_{n-1}}{D_{n-1}} = \frac{y''_{n-1} - y''_{n-2}}{D_{n-2}}
\]

or \(D_1y''_3 - (D_1 + D_2)y''_2 + D_2y''_1 = 0\)

and \(D_{n-2}y''_n - (D_{n-2} + D_{n-1})y''_{n-1} + D_{n-1}y''_{n-2} = 0\).

The matrix to be inverted to obtain \(y''\) is almost tri-diagonal, which is straightforward. Had we used the rounded uniform cubic spline with large \(\alpha\) this matrix would have been triangular which is then more difficult to invert. However, mathematically speaking the results are identical. We have used a general library routine LINSYS to solve the above linear equations. Having obtained values of \(y''\), we find \(y'\) from eq. (3). Interpolation was then performed using eq. (1) for a given number of points between the known reference points.

Chebyshev polynomials were subsequently used to obtain a smooth fit of the model. For this we used the routine CHCOF1 which again calls the interpolation routine eq. (1). Tests were made to find to what extent the orthogonal Chebyshev coefficients were significant. We consider a coefficient, \(c_i\), insignificant if

\[|c_i| < \epsilon \quad \text{and} \quad |c_{i-2}| < \epsilon,\]

where \(\epsilon\) is an estimate of the error. The reason for this is that one can expect the curve to be dominantly asymmetric. Therefore, we determined the number of significant coefficients in each symmetry class separately. Insignificant coefficients were set to zero to obtain an optimum smooth fit. The relative accuracy chosen was \(2 \times 10^{-5}\).

Plots of the interpolated values, \(y'\) and \(y''\) are made to check the results and for comparison of the exact and smooth fits (figs 2-7). We also make a plot of the differences of the smoothed fit and the original observations (fig. 8).
2. PROGRAM DESCRIPTION

2.1 Introduction

The program fits calibration curves to Hall probes by using a regular cubic spline model as a method of interpolation for an exact fit and Chebyshev polynomials for a smooth fit.

2.2 Description of variables and constants

Y  
are actual field values (y)

WORK 
is a working array which is a parameter for library routine LINSYS

A  
is the matrix to be solved by LINSYS

YPR 
are the function values of y

X  
are the actual Hall voltages (x)

YP  
are the y'-values computed

YPP  
are the y''-values computed

CHC  
are the Chebyshev coefficients returned by the library routine CHCOFL

R  
are the right-hand sides of the equations to be solved and are passed over to library routine LINSYS

YS  
are the differences between the exact and smooth values at the given reference points

F1  
are the lowest field values for each probe

F2  
are the highest field values for each probe

XN  
are the x-values for one probe for a given range of field values in ascending order of magnitude

YN  
are the y-values for one probe for a given range of field values in ascending order of magnitude

D  
are the interval sizes of XN (D(I) = XN(I+1) - XN(I))
YPN  are \( y' \)-values computed from Taylor series having obtained \( y'' \) from library routine LINSYS

YPPN  are \( y'' \)-values returned by routine LINSYS

NDATA  are the actual number of reference points considered

S  is an indication of order of \( x \)-values: +1 for ascending order and -1 for descending order

MNEW  is the array index of the last Chebyshev coefficient not set to zero.

IPROBE  indicates the current probe number

I1  is the index value of the lowest field value specified

I2  is the index value of the highest field value specified

XMIN  is the minimum \( x \)-value

XMAX  is the maximum \( x \)-value

XX  is the current interpolated value of \( x \)

DX  is the interval size of the interpolated values of \( x \)

M  is the number of Chebyshev coefficients returned by CHCOF1

XXX  is the adjusted value of \( xx \), passed as a parameter to CHSUM1, normalized to the interval \((-1,1)\)

MDIM1  is the dimension of \( Y,A,X,R,YS,XN,YN,D,YPN \) and YPPN (set to 80)

MDIM2  is the dimension of \( YPR,YP,YPP \) and XPLT (set to 500)

MDIM3  is the dimension of \( X, F1 \) and \( F2 \) (set to 12)

SQ  is the sum of squares of the residuals at observed points

NP  is the number of Hall probes

N  is the number of reference points per set of Hall probes

NPR  is the number of interpolated values to be calculated and plotted

ERRORU  is the lower level of relative accuracy

ERRORL  is the upper level of relative accuracy.
2.3 Description of input

Card 1  NP, N, NPR, ERRORU, ERRORL - format 3I5,2E10.3
Card 2  F1(1), F2(2) - format 2F10.0
        :  :  :  :
Card NP+1 F1(NP), F2(NP)

where array F1 holds lowest field values for each probe and array F2 holds
highest field values for each probe.

Card NP+2 onwards are the data cards.

    Each card has x-values for 4 probes and y-value - format 5F10.0.

Nota:  The format should be changed where number of probes is not 4.
       NPR should be less than or equal to zero if no plots are desired
       (see also note (i) page 11).

2.4 Description of output

Line printer output

(a) NP,N,NPR,ERRORU,ERRORL are output - format 3I5,2E10.3, where ERRORU
    and/or ERRORL are input as less than or equal to zero, the program-
    assigned values are printed.

(b) The x-values for each probe and y-values are printed - format 5E15.7.

(c) The sum of squares of residuals from the exact fit at observed points
    is printed - format E20.10.

(d) The current probe number, minimum and maximum x-values and number
    of calibration coefficients are printed - format I10,2E15.7,I10.

(e) The calibration coefficients for the current probe are printed -
    format 6E13.6.

(f) A list of coefficients set to zero is printed giving the previous
    values - format E13.7.

(g) The index value of the last non-zero coefficient is printed -
    format I5.
(h) A lineprinter plot for each (independent of the value of NPR) is made of the differences between the smoothed fit and the original observations.

2.5 Error messages and recovery

2.5.1 NPR greater than MDIM2

The following message is output:

'User Error-NPR set to VALUE OF NPR
Abort: dimensions only allow NPR up to VALUE OF MDIM2'
Recovery - the user should check the value of NPR and, if correct, change the dimensions of XPL0T,YPR,YP and YPP accordingly.

2.5.2 N greater than MDIM1

The following message is output:

'User error-N set to VALUE OF N
Abort: dimensions only allow N up to VALUE OF MDIM1'
Recovery - the user should check the value of N and, if correct, change the dimensions of Y,A,X,R,YS,XN,YN,D,YPN and YPPN.

2.5.3 NP greater than MDIM3

The following message is output:

'User error-NP set to VALUE OF NP
Abort: dimensions only allow NP up to VALUE OF NP'
Recovery - the user should check the value of NP and, if correct, change the dimension of X, F1 and F2.

2.5.4 F1(IPROBE) = F2(IPROBE)

The following message is output:

'User error - maximum and minimum field values equal. F1 = VALUE OF F1(IPROBE), F2 = VALUE OF F2(IPROBE) abort'
Recovery - the user has inadvertently set the maximum and minimum field values to same value. Check input values.
2.6 Punch card output

(a) Current probe number, minimum and maximum x-values and number of
calibration coefficients are punched on cards - format 110.2E15.7,110.

(b) The calibration coefficients for the current probe are punched,
6 values per card - format 6E13.6.

2.7 Graph plots

Where NPR is greater than zero, the following plots of the exact and
smooth fits are made for each probe via the graph plotter:

(i) fitted values of y
(ii) fitted values of y'
(iii) fitted values of y"
(iv) differences between the smoothed fit and the
original observations.

2.8 Algorithm for program TSPLWN

Read NP,N,NPR,ERRORU,ERRORL
Read F1 and F2 for each probe
If NPR greater than MDIM2 output message and jump to END
If N greater than MDIM1 output message and jump to END
If NP greater than MDIM3 output message and jump to END
If ERRORU and/or ERRORL less than or equal to 1.E-20 set to
appropriate value(s).
Output NP,N,NPR,ERRORU,ERRORL
Read x and y-values for each probe
Print x and y-values for each probe
REPEAT until all probes handled

Calculate index of lowest (I1) and highest (I2) reference points by
comparing y-values with ERRORU x F2 for particular probe
Calculate number of reference points to be used (NDATA)

Get minimum Hall voltage (XMIN) and maximum Hall voltage (XMAX) for this probe

Determine ordering of Hall voltages and put Hall voltages and field values into ascending order of magnitude, if necessary

Calculate differences between consecutive voltages \((D_i = x_{i+1} - x_i)\) for NDATA points

Get right-hand sides of equations

Set up matrix \(A\) with zero elements

Set non-zero elements of \(A\) by imposing continuity constraints and marking third derivative zero at second and last but one points

Call function LINSYS to solve matrix \(A\) returning second derivative \((y''_i)\)

Calculate first derivative \((y'_i)\)

If no plots desired, jump to \(a\)

Calculate interval size (DX) for NPR points

Calculate exact spline fit by calculating \(y\)-value (YPR) for each interpolated value (XX) by calling routine FUNT

Make plot (via graph plotter) of YPR using library plotting routines (TVBGN, TVRNG, TVGRPH, TVTEXT, TVNUMA, TVEND)

Calculate \(y'_i\) and make plot (via graph plotter) of \(y'_i\) using library routines

Calculate \(y''_i\) and make plot (via graph plotter) of \(y''_i\)-values

\((a)\) Evaluate and output to lineprinter sum of squares of residuals at the observed points

Call library routine CHCOF1 to calculate Chebyshev coefficients

Output to lineprinter and punch cards IPROBE, XMIN, XMAX and M
Output Chebyshev coefficients to lineprinter and punch cards
Set insignificant coefficients to zero and output those set to zero
with their previous values
Output index of last coefficient not set to zero
If no plots desired jump to (b)

Call library routine CHSUM1 to get Chebyshev expansion of interpolated
points (YPR)
Calculate $y_i'$ of smooth fit
Make plots (via graph plotter) of YPR and $y_i'$ of smooth fit
Calculate $y_i''$ of smooth fit and make plot (via graph plotter)

(b) Calculate differences between the original values and those obtained
from smooth fit (YS)
If no plots desired, jump to (c)
Calculate maximum and minimum values of differences (YS) for plotting
routine
Plot differences (YS) via graph plotter

(c) Plot differences on lineprinter using program subroutine (PLOT)

2.9 Program access

The program, TSPLWN, is catalogued on the MFB machine as FITCALCURVES
and may be accessed under ID = EP078WIND

2.10 To run the program

The following card deck should be used:

1. GGNNN,T14.  (GG is the delivery code for the appropriate RIOS and
NNN are any alphanumerics)
2. ACCOUNT,WIND,EP,007870.
3. FIND,LIB1,7600LIBRARY,ID=PROGLIB.
4. FIND, MAN, MAN, ID=OPS SYSTEM.
5. LIBRARY, LIB1, MAN.
7. ATTACH, FITCAL, FITCALCURVES, ID=EP078WIND, ST=CCQ.
8. FILE, TAPE4, RT=W, BT=I.
9. COPY, INPUT, TAPE5.
10. REWIND, TAPE5.
11. FTN, L, R, U, I=FITCAL.
12. LDSET, MAP=B/ZZZMP.
13. LGO.
14. DISPOSE, OUTPUT, *PR, ST=CCPIrr. (rr is the identifier for the appropriate RIOS printer)
15. CATALOG, TAPE4, PLOTS, ID=EP078WIND, ST=CCQ.
16. PLOTLIM(80).
17. CPPILOT, TAPE4.
18. DISPOSE, PLOT, *PT=CS1, ST=CCP.

End of record card.

The data cards should follow in the format specified under "Description of input" sect. 2.3.

Notes

(i) If NPR has been set less than or equal to zero (i.e. no plots are wanted), then control cards 8, 15, 16, 17 and 18 are not required.

(ii) The plots, catalogued under PLOTS (14 day limit), may be viewed at a terminal under Intercom. The following commands should be used:

ETL, 300
ATTACH, CERNJOB
CERNJOB, TV=PLOTS, ID=EP78WIND

The prompt TVCMD= is then output to the screen. The user may then enter D to obtain a list of the available commands. The plots may then be viewed using these commands. Example: EBFO
clears the screen, and outputs frame 0 (i.e. first plot) in a rectangular region. Further information on the graphics facilities may be found in "GD3 Introductory Guide" - C.S. Curran and J.M. Howie, CERN/DD/US/8 (25 May 1977).

(iii) Control card 16 gives a plotting time limit on the graph plotter of 80 mins. This is adequate for the plots produced for up to 4 probes. The time limit should be increased accordingly for a greater number of probes.

(iv) The punched card output and plots (where specified) will be delivered to the RIOS specified on the JOB card. The lineprinter output will be printed at the RIOS specified on the DISPOSE card (control card 14).

REFERENCE
