# INHERENT FLEXIBILITY OF MATHEMATICAL SPLINES FOR PULSE SHAPE RECONSTRUCTION


## Aba A. Bentil Andam,

PhD, CPhys, MInstP
Department of Physics,
Universty of Science & Technology, Kumasi, Ghana



## ABSTRACT

Effective retrieval of information from digitised pulse shape data depends on a good data fitting procedure. Polynomial approximation, although easy to compute, results in curve oscillation and interpolation when high order polynomials are used. In addition, the analyticity of polynomials introduces a major drawback of global dependence on local properties in the interval of approximation.

The use of mathematical splines helps to overcome these problems. A quartic spline of order 5 with 6 knots has been used to reconstruct digitised Cerenkov light pulse shapes. The pulse shape parameters measured are consistent with measurements from other experiments.


*Keywords: pulse shape, data fitting, minimization, analyticity, spline, knots.*

## INTRODUCTION

The backbone of any information retrieval from digitised pulse shape data is an effective data fitting procedure. The fitted curve should not merely be smooth for aesthetic satisfaction but should permit the efficient computation of the parameters of the original pulse.

For a given set of histogram points

$$x_1 < x_2 < \dots < x_n < x_{n+1}$$

corresponding to heights

$$h_1, h_2, \dots \quad \dots \quad , h_n$$

where $h_i$ is the height of the histogram over the interval $[x_i, x_{i+1}]$, (Fig. 1), a fitting function, g (x), and an appropriate data fitting procedure must be chosen to satisfy the least squares minimization condition

$$\mathcal{E}^2 = \left[ \int_{x_i}^{x_{n+1}} g(x)dx - \sum_{i=1}^{n} h_i \; \Delta x_i \right]^2 \longrightarrow 0 \qquad (1)$$

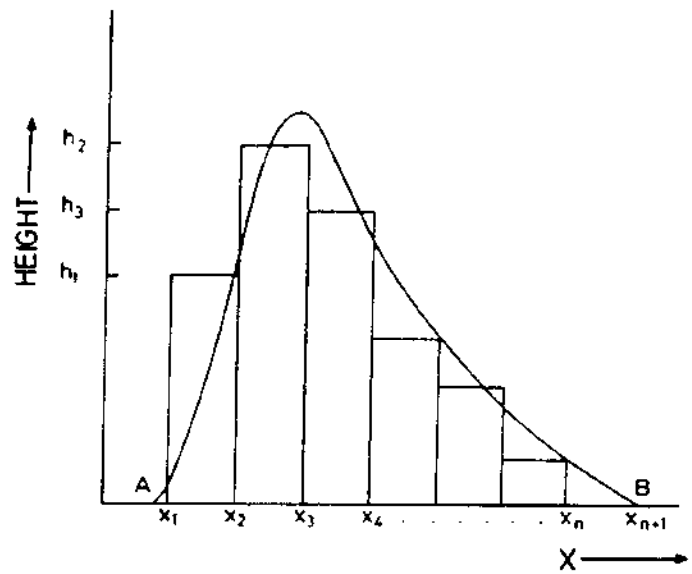


Figure 1: "Area Matching" Curve Approximation To a Histogram

## POLYNOMIAL APPROXIMATION

As a first approximation, the space $G_n$ of polynomials of order m may be considered, where

$$\mathcal{G}_n = \left\{ g(x) : g(x) = \sum_{i=1}^{n} C_i x^{i-1}, C_1, \dots, C_n, \; x \text{ real} \right\} \qquad (2)$$

The popularity of polynomials for approximation stems from the ease of computation with the basic arithmetic operations. However, serious smoothing problems are encountered with high order polynomials, resulting in curve oscillation and interpolation at the cost of smoothing. In addition, the analyticity of polynomials introduces a major drawback of global dependence on local properties in the interval of approximation.

Most physical quantities have functional forms whose behaviour in one region of the approximation interval does not necessarily reflect their behaviour throughout the interval. Any fitting procedure to practical data must therefore be able to allow for such variations in functional form.

## PEARSON TYPE III FUNCTION

Many pulse shapes may be fitted with a unimodal fast-rising, slow-decaying function. One class of functions which satisfy this condition, and are quick to compute, is the Pearson type III function, of general form:

$$P(t) = A(t+\tau)^{\delta} e^{-\left\{-(t+\tau)/\beta\right\}} \qquad (3)$$

where $\beta$, $\delta$, and $\tau$ are constants, with A as a normalisation factor.

By combining different values of $\beta$ and $\delta$, distributions with varying degrees of skewness and kurtosis may be produced (Fig.2). A useful approximation which saves computer time is

$$\int_0^\infty x^n e^{-ax}\, dx = \frac{\Gamma(n+1)}{a^{n+1}} \qquad (4)$$

Hence,

$$\int_0^\infty P(t)dt = A\left[\frac{a^{n+1}}{\Gamma(n+1)}\right]\int_0^\infty \left(t^n e^{-ax}\right) dt \qquad (5)$$

where

$n = \delta$

$a = 1/\beta$

and A = area of the pulse, used as a normalization factor.

Computer simulated pulses of different widths and sizes have been fitted with the Pearson type III function. The results, summarized in Fig.3, indicate that Pearson functions can cope with the reconstruction of Cerenkov light pulses of height >50mV and medium width ($\gtrsim$ 20ns). For narrow pulses there is a systematic over-estimate of the pulse width and under-estimate of the height. This might be due to the strong effect of the exponential part of eqn.5, which causes the fitted shape to fall before it has developed along the full length of the leading edges of the pulse.

In order to cover a wide spectrum of pulse sizes and widths expected from a large experimental database, it is necessary to use a fitting function that is more flexible than ordinary polynomials or Pearson-type functions, such as mathematical splines.

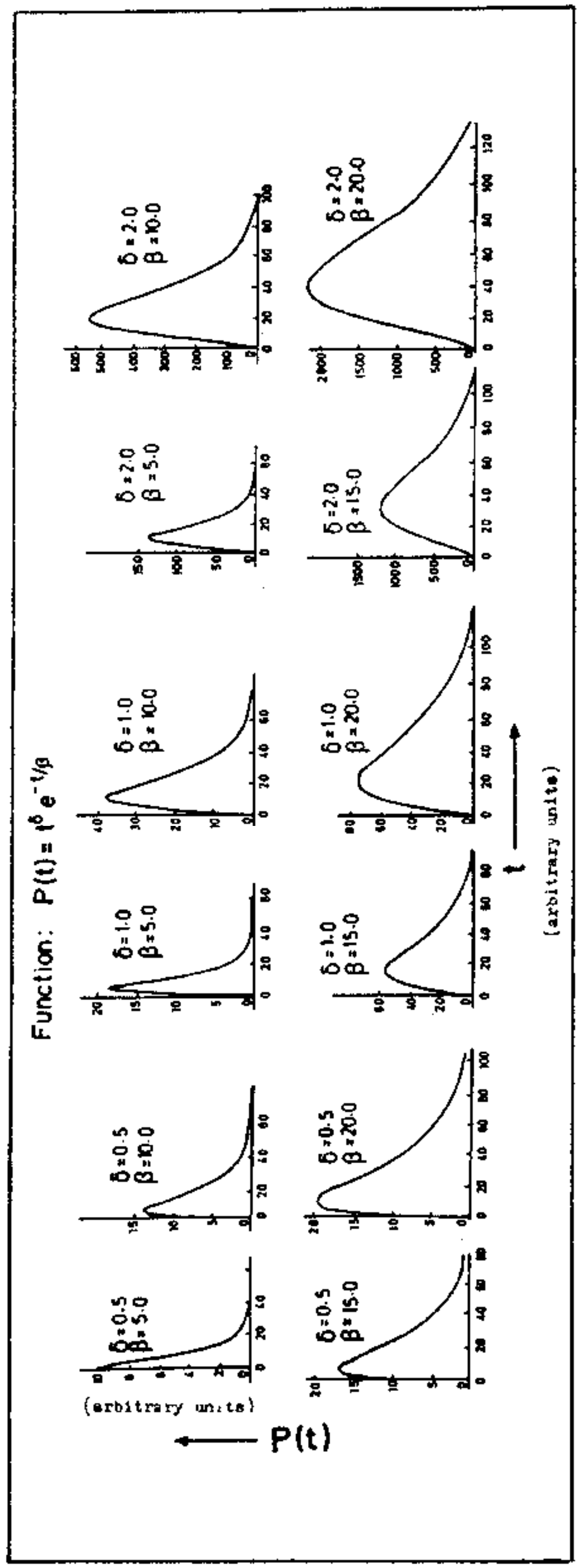


Figure 2: Some of the Different Distributions Obtained by Various Combinations of $\delta$ and $\beta$ In the Pearson Function: $P(t) = t^{\delta} e^{-t/\beta}$

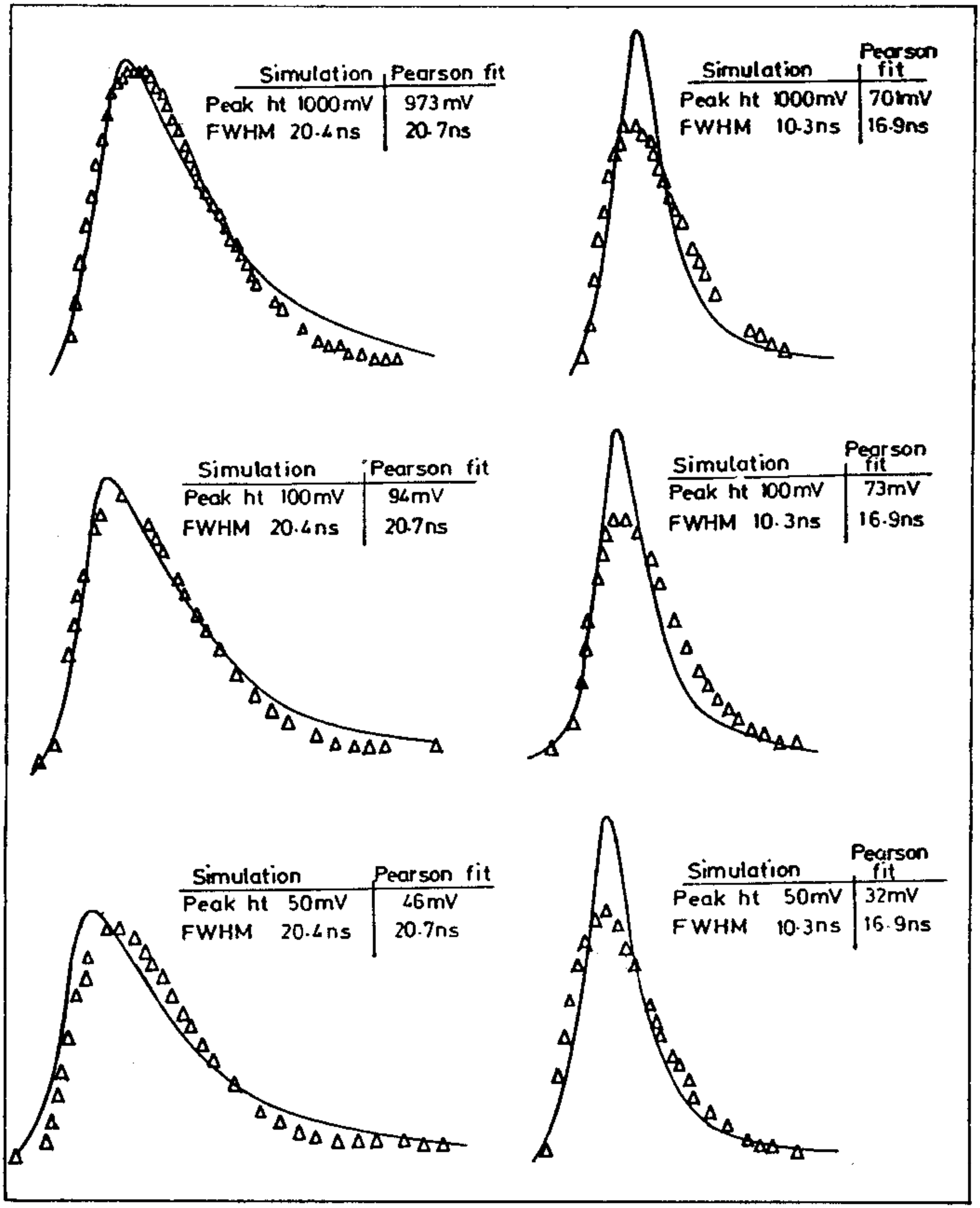


Figure 3: Pearson Function Fits to Simulated Cerenkov Light Pulses.
— Input Simulated Pulse △ Pearson Fit

Pearson-type functions, such as mathematical splines.

## MATHEMATICAL SPLINES

Mathematical splines are the conceptual equivalent of the long-known draughtsmen's tool consisting of a strip of wood or other elastic material, anchored in place with lead weights at given points. The strip of wood, or "spline" is used to fair in a smooth curve between a set of points by varying the position of the weights and the spline. In recent years, the traditional spline has been replaced by the Flexi-curve.

If the draughtsman's spline is considered as a thin beam, then its bending moment, M(X), and Young's modulus, E, will satisfy the Bernoulli-Euler Law.

$$M(x) = EI\,[1/R(x)] \qquad (6)$$

where I is the geometric moment of inertia and R(x) is the radius of curvature of the curve assumed by the deformed axis of the beam, i.e. the elastica.

By replacing the mechanical spline with its elastica, mathematical splines may be formulated as localised fits or piecewise polynomials whose segments are defined only in a limited range of the independent variable, with the constraint that the polynomial segments must have continuity of function and derivative at the joints or "knots". This inherent property gives rise to a function that is smooth and continuous anywhere inside the boundary knots, but vanishes outside the boundaries.

For many important applications in the Pure and Applied Sciences this mathematical model of the draughtsman's spline has been found to be adequate and realistic, e.g. in approximation theory, numerical analysis, and data fitting.

Consider a set of real numbers, strictly increasing in the order

$$x_1, x_2, \ldots \quad \ldots, x_n$$

We define a spline function, $S(x)$, of degree m (or order m + 1) with the knots

$$x_1, x_2 \ldots \quad \ldots, x_n$$

as a function defined on the entire real line $x_1 \ldots \ldots x_n$ and having the following properties:

i) In each interval $(x_i, x_{i+1})$ where

$$i = 0, 1, \ldots \quad \ldots, n,$$

$S(x)$ is given by some polynomial of degree m or less.

ii) $S(x)$ and its derivatives of order

$$1, 2, \ldots\ldots\ldots, m-1$$

are continuous everywhere in the interval of approximation.

Consequently, the spline may be divided into sections or separate entities, provided the constraints of continuity of function, f(x), slope, f (x), and curvature, f (x), are met for adjacent polynomials, and the spline function chosen satisfies the minimization condition in equation(1).

For example, the Taylor series may be used as the fitting function, but equally suitable functions include the exponential functions, whose limit is the Gaussian. Among the simplest spline functions is a cubic "basic spline curve" or B-spline, [1], [2], whose basis is the truncated power function:

$$x_+^n = \begin{cases} x^n, & \text{for } x \geq 0 \\ 0, & \text{for } x < 0 \end{cases} \qquad (7)$$

Definition

Let

$$F(x) = (x_r \quad - \quad x)_+^{n-1}$$

$$= \begin{cases} (x_r - x)^{n-1}, & \text{for } x_r \geq x \\ 0, & \text{for } x_r < x \end{cases} \qquad (8)$$

Then the B-spline can be evaluated on the r^th knot position, $x_r$ using equation (7), and the divided differences of $F(x)$ in $x_r$ for any fixed $x$.

Let $M(x_o, x_1)$ be the first divided difference of $F(x)$ on $x_o$ and $x_1$

$$M(x_o, x_1) = \frac{F(x_1) - F(x_o)}{x_1 - x_o} \qquad (9)$$

The divided diference of order m is

$$M(x_j; x_o; \ldots, x_m) = \frac{M(x_j; x_1, \ldots, x_m) - M(x_j; x_o, \ldots X_{m-1})}{x_m - x_o} \qquad (10)$$

From which it follows that

$$M(x_j; x_o, x_1) = \frac{M(x_j, x_1) - M(x_j, x_o)}{x_1 - x_o} \qquad (11)$$

and

$$M(x_j, x_j) = F(x_j) \qquad (12)$$

12. Chantler M.P., Craig M.A.B., McComb T.J.L., Orford K.J., Turver K.E. and Wallby G.M.; Proc. 17th Int.Conf.on Cosmic Rays, Paris 6(1981), 121

13. Hammond R.T., Orford K.J., Protheroe R.J. Shearer J.A.L., Turver K.E., Waddoup W.D. and Wellby D.W., Nouvo Cimento 1C (1978), 315

The B-spline, $M_{ni}$, of order n (or degree n-1) is then defined as;

$$M_{ni}(x) = M(x_i; x_o, x_1, \ldots, x_n)$$

$$= \sum_{i=o}^{n} \frac{F(x)}{\prod_{j=o,\ j \neq 1}^{n} (x_i - x_j)} \qquad (13)$$

For computational purposes, the B-spline may be normalized by utilising the property:

$$\int_{-\infty}^{\infty} M_{ni}(x)\, dx = {}^1/_n \qquad (14)$$

## ALGORITHM

Although B-splines may be evaluated directly from the divided differences definition, Cox [3] has pointed out that some evaluation may fail because of cancellation of nearly equal terms, and proposed a stable method based on the recurrence relation:

$$M_{ni}(x) = \frac{(x - x_{i-n})\, M_{n-1,i-1}(x) + (x_i - x)\, M_{n-1,i}(x)}{x_i - x_{i-n}} \qquad (15)$$

where $M_{ni}$ is a spline of order n ending on knot $x_i$; $M_{n-1,i}(x)$, $M_{n-1,i}(x)$ are $(n-1)$th divided difference on $x_{i-1}$ and $x_i$ respectively

In pulse shape analysis there is a physical constraint of fitting a unimodal spline with 2 fixed end knots, because of the finite duration of a light pulse.

A useful relationship which may be used to simplify computation is

$$M_{ij}(x) = \begin{cases} 1/(x_j - x_{j-1}), & \text{for } x_{j-1} \leq x \\ 0, & \text{otherwise} \end{cases} \qquad (16)$$

Hence, for $x_{i-2} \leqslant x \leqslant x_{i-1}$, all the terms in the first divided difference column are zero except $M_i$, $i-1$. The computation is then reduced to only the terms in the rhomboidal array in Table 1.

**TABLE 1**
**RHOMBOIDAL ARRAY OF ELEMENTS FOR A QUARTIC SPLINE OBTAINED BY APPLYING THE RELATION IN EQUATION (16)**

| | 1st divided differences | 2nd divided differences | 3rd divided differences | 4thdivided differences | 5thdivided differences |
|---|---|---|---|---|---|
| $x_{i-5}$ | | | | | |
| | 0 | | | | |
| $x_{i-4}$ | | 0 | | | |
| | 0 | | 0 | | |
| $x_{i-3}$ | | 0 | | $M_{4,i-1}$ | |
| | 0 | | $M_{3,i-1}$ | | $M_{5,i}$ |
| $x_{i-2}$ | | $M_{2,i-1}$ | | $M_{4,i}$ | |
| | $M_{1,i-1}$ | | $M_{3,i}$ | | |
| $x_{i-1}$ | | $M_{2,i}$ | | | |
| | 0 | | | | |
| $x_i$ | | | | | |

## FIXED OR VARIABLE KNOT POSITIONS?

De Boor and Rice [4,5] and Schumaker [6,7] have shown that approximation to data by splines improves greatly if the knots are free variables.

In the present work, quartic splines were computed with variable-knot positions. The knot positions were allowed to vary between pre-set boundaries, while satisfying non-linear least squares minimisation criteria, by the use of MINUIT [8].

By combining the Monte Carlo searching procedure, SEEK, with minimisation techniques, MINUIT incorporates the SIMPLEX method of Nelder and Mead [9] and MIGRAD [10], to obtain optimal knot locations which are then used for computing the spline.

## WEIGHTED SPLINES

One refinement in analysing digitised pulses, in particular pulses for which the presence of undershoot can result in inaccurate fits, is to fit a number of splines beginning and ending on the knots, and take a weighted spline over the entire interval (see Figure 4).

For this purpose the number of knots will have to be increased beyond the range of the data. Hence in Figure 4 the knots $T_{i-1}$, $T_{i-2}$, $T_{i-3}$ are incorporated before the first knot $T_i$ of the data, while the knots $T_{i+4}$, $T_{i+5}$, $T_{i+6}$ are used after the last knot of the data. Only part of each B-spline whose knots go beyond the data boundary will contribute to the final spline. These sub-splines will have less weighting than the sub-splines which span the data set.

The resultant spline (S(t) is therefore given by:

$$S(t) = C_i M_{4i}(t) + C_{i+1} M_{4i+1}(t) + \dots \dots + C_{i+n} M_{4i+n}(t) \qquad (17)$$

where

$$M_{4i}(t), M_{4i+1}(t), \dots \dots M_{4i+n}(t)$$

are the splines beginning and ending on the knots, and

$$C_i, C_{i+1}, \dots \dots C_{i+n}$$

are the coefficients appropriate to each spline.

Hence, each sub-spline would be weighted ac-

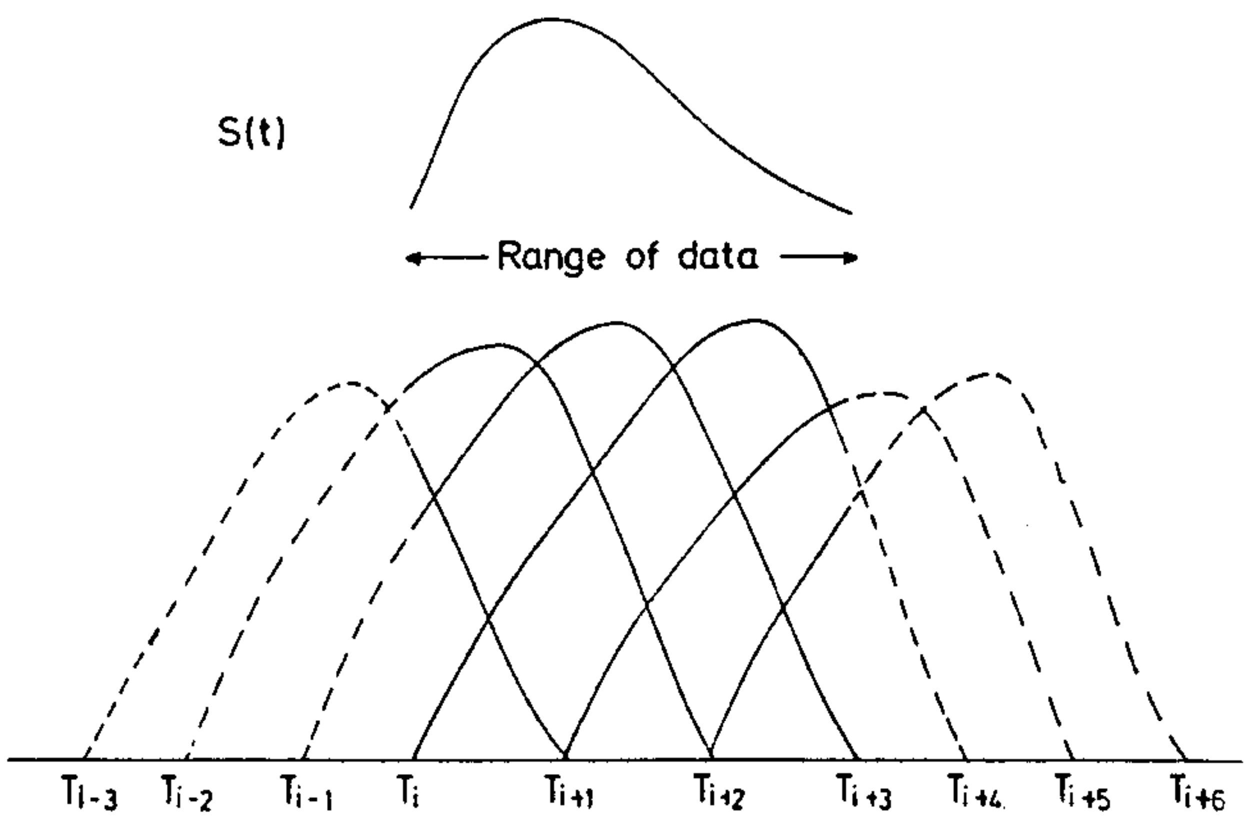


Figure 4: Weighted B-Splines for Non-Linear Least Squares Fit

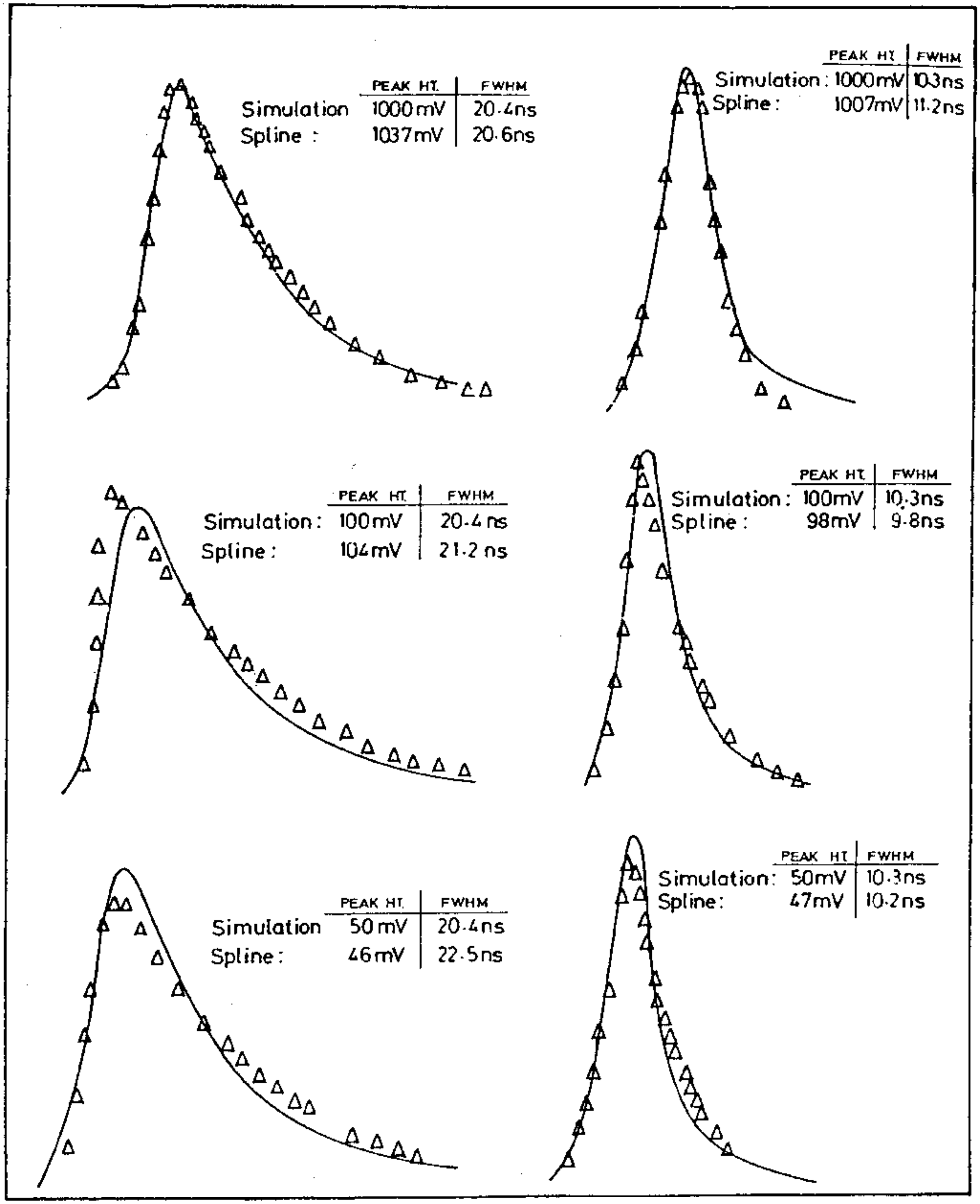


Figure 5: Quartic Spline Fits to Simulated Cerekov Light Pulses
—— Input Simulated Pulse △ Spline Fits

cording to the sections of the pulse it covers. Equation (17) may be solved by the matrix equation:

$$\begin{pmatrix} M_1(t_1) & M_2(t_1) \dots & M_n(t_1) \\ M_1(t_2) & \dots & M_n(t_2) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ M_1(t_{n-1}) & \dots & M_n(t_{n-1}) \\ M_1(t_n) & \dots & M_n(t_n) \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ \cdot \\ \cdot \\ \cdot \\ C_{n-1} \\ C_n \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \\ \cdot \\ \cdot \\ \cdot \\ V_{n-1} \\ V_n \end{pmatrix} \qquad (18)$$

$$\text{or} \quad \underset{\sim}{M} \times \underset{\sim}{C} = \underset{\sim}{V} \qquad (19)$$

where

M is a function of the knot position only.

## QUARTIC SPLINE FITS TO SIMULATED AND EXPERIMENTAL DATA

Figure 5 shows a set of computer simulated light pulses which were spline-fitted. A close approximation was obtained between the fitted splines and the simulated data over all pulse sizes tested. The fitting function used was a quartic spline with 6 knots as described earlier in this work.

The experimental data were obtained from the Dugway Extensive Air Shower Experiment [11,12]. An array of eight detectors was used to record the shape of atmospheric Cerenkov light pulse from extensive air showers (EAS) of primary energy $10^{16}$ - $10^{18}$ eV. The light pulse was measured sequentially as charge, digitised in narrow time intervals ("slices") at each detector. The amount of charge in each time interval of 10ns was recorded if the amplified signal from the detector exceeded a preset discriminator threshold of 20 mV.

Information for one pulse comprised a maximum of 7 (data) points (i.e. a maximum of 6 "slices" plus 1 discriminator level). The constraint of 2 end knots leaves 5 degrees of freedom available. A quartic spline with n = 5 and 6 knots was chosen as an appropriate fitting function to the data.

## CONCLUSION

The validity of the spline-fitting procedure is evident in the extensive air shower (EAS) measurements derived from the computed pulse shape parameters [11]. The pulse Peak Height, Rise Time and full width at half maximum (FWHM) are all accurately determined from the fits.

The derived extensive air shower characteristics reported in [11] were found to be consistent with EAS measuremens derived from other Cerenkov light parameters, such as the lateral distribution of the Cerenkov light [12], and from experiments using different techniques (e.g. [13]). In particular, the depth of electron cascade maximum of the extensive air shower determined from the FWHM measurements [14] helped to augment data on cosmic radiation.

The results obtained in this work indicate that spline-fitting is adequate for reconstruction of digitised Cerenkov light pulses. The Peak Height, Rise Time, and full width athalf maximum (FWHM) are all determined acurately from the fitting procedure. Extensive Air Shower (EAS) measurements computed from these pulse shape parameters are consistent with measurements from other experiments.

## ACKNOWLEDGEMENTS

This paper was prepared while the author was on an Associate visit to the International Centre for Theoretical Physics (ICTP), Trieste, Italy. It is a pleasure to thank the ICTP for use of their facilities, and the Swedish Agency for Research Co-operation with Developing Countries (SAREC), for their financial support of the ICTP Associate Scheme.

### *REFERENCES*

*1. Schoenberg, I.J; "Contributions to the problem of Approximation of Equidistant Data by Analysis Functions" Quart.Appl. Maths. 4, (1946), 45-99, 122-141.*

*2. Schoenberg, I.J,; "On Spline Functions" in Inequalities II, O. Shisha, Ed. Academic Press, NY(1967) .335-374*

*3. Cox M.G.; J.Inst. Maths Applics. 10, (1972), 134*

*4. de Boor C. and Rice J.R.; "Least Squares Cubic Spline Approximation I: Fixed Knots" CST TR 20, Purdue University, 1968.*

*5. de Boor C and Rice J.R.; "Least Squares Cubic Spline Approximation II: Variable Knots" CSD TR 21, Purdue University, 1968.*

*6. Schumaker L.L.; "Some Algorithms for Splines" in Theory and Applications of Spline Functions. T.N.E. Greville ed., Academic Press, (1969),pp.87-102.*

*7. Schumaker L.L.; "Spline Functions: Basic Theory", John Wiley & Sons, 1981*

*8. James M. and Roos M.; Comp.Phys. Comm.10,(1967),343*

*9. Nelder J.A. and Mead R.; Computer Journal,7(1967), 308*

*10. Fletcher R.; Computer Journal 13 (1970), 317*

*11. Andam A.A., Chantler M.P. Craig M.A.B. McComb T.J.L., Orford K.J., Turver K.E., and Walley G.M.., Proc.17th Int.Conf. on Cosmic Rays, Paris, 6(1981), 125*