Internal alignment of the SLD vertex detector

C.D. Jackson[a], D. Su[b], F.J. Wickens[a]
for the SLD Collaboration

[a] *Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, UK*
[b] *Stanford Linear Accelerator Center, Stanford, CA 94309, USA*

Abstract
The tracking resolution and vertex finding capabilities of the SLD experiment depended upon a precise knowledge of the location and orientation of the elements of the SLD pixel vertex detector (VXD3) in 3D space. At the heart of the procedure described here to align the 96 CCDs is the matrix inversion technique of singular value decomposition (SVD). This tool was employed to unfold the detector geometry corrections from the track data in the VXD3. The algorithm was adapted to perform an optimal χ^2 minimization by careful treatment of the track hit residual measurement errors. The tracking resolution obtained with the aligned geometry achieved the design performance. Comments are given on how this method could be used for other trackers.

7.1 Introduction

Between 1996 and 1998 around 400 000 hadronic Z^0 decays were recorded by the SLC Large Detector (SLD) at the e^+e^- SLAC Linear Collider (SLC). The long lifetimes of B and D hadrons results in tracks from $Z^0 \to b\bar{b}$ and $Z^0 \to c\bar{c}$ events having more than one origin: fragmentation tracks form the primary vertex at the e^+e^- Interaction Point (IP) while tracks from the decays of the heavy hadrons form displaced secondary vertices.

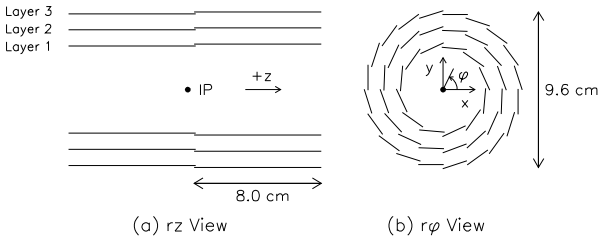

Fig. 7.1: Layout of the CCDs in VXD3: (a) cross-section through the rz plane, (b) view from $+z$ end with each line representing one ladder supporting two CCDs

An important part of the physics programme at SLD relied on the identification of these secondary vertices [1]. The ability to do this depended on the tracking resolution, and particularly the alignment of the vertex detector elements. This paper describes the technique developed for the internal alignment of the components of the SLD vertex detector which allowed the design performance to be achieved.

The SLD Vertex Detector (VXD3 [2]) consisted of 96 charge-coupled devices (CCDs [3]) arranged on 3 cylindrical layers of supporting ladders, with radii of 2.7, 3.7 and 4.7 cm, around the IP as shown in Fig. 7.1. The z-axis was parallel to the beam-line with the positive direction towards the North face of the SLD detector. There were 48 CCDs facing the beam on the South half of each ladder together with 48 outward facing CCDs on the North half, with an overlap of around 1 mm, called a 'Doublet', in the region near $z = 0$. There was an overlap of similar size between neighbouring CCDs in the same layer, known as a 'Shingle', in the $r\phi$ plane, Fig. 7.1(b). The design of the VXD3, installed in 1996, was motivated by the experience with the previous SLD vertex detector VXD2. In particular, the Doublet and Shingle overlaps are a crucial feature that not only secure the coverage and avoids gaps, but were also introduced with tracking based alignment in mind. Each of the 96 CCDs measures 8 cm in length by 1.6 cm width and houses 3.2 million 20×20 μm pixels. A charged track traversing a CCD leaves a hit consisting of a cluster of several pixels, within each of which the ionization charge was recorded. The centroid of the cluster determines a point on the track trajectory with an intrinsic resolution estimated to be around 3.5 μm.

Sets of such hits on different CCDs within small

solid angular regions about the IP were assumed to have been produced by the passing of a single charged track. Charged tracks from the interaction region were identified and reconstructed initially in the Central Drift Chamber (CDC) [4]. The helix parameters of each track, in the 0.6 T solenoid field, were used to determine the region of the VXD3 traversed by the corresponding charged particle and to associate the track with the VXD3 hits. Since the VXD3 hit efficiency was very high most tracks that pass through all three layers of the detector were found to associate with hits on three CCDs (because of overlaps some tracks were seen on more than three CCDs).

In a preamble to the full internal alignment the CCD geometry was corrected for deviations from the ideal geometry location and a flat surface with measurements from a room temperature optical survey. The calculated effects of gravitational sag were also taken into account. Further, relying both on predicted effects and observation of detector performance, the CCD data were corrected for several shifted and smeared electronic channels in the 1996 run and for mechanical contraction inside the 200 K VXD3 cryostat (lowered again to 180 K before the 1997 run). These initial corrections determine the 'nominal' geometry of the detector, with typically a few tens of microns precision, with which the first data was reconstructed. Subsequently a χ^2 was calculated to measure the deviation of the linked VXD3 hit from the CDC track trajectory taking into account the tracking resolution. Initially the VXD3 is treated as a rigid body and globally aligned to the CDC by minimizing the $\sum \chi^2$ over tracks in a large number of events with respect to the three translational and three rotational degrees of freedom of the whole VXD3. The impact parameter resolution at the IP of tracks fitted to the combined CDC and VXD3 hit data then depended primarily upon the CCD single hit resolution. This in turn was the product of the intrinsic resolution and the systematic uncertainties in the relative internal locations and orientations of each of the 96 CCDs with respect to the same set of six degrees of freedom. The aim of the internal alignment was to remove the latter contribution with the determination of these $96 \times 6 = 576$ geometric corrections. Before the physics analysis could fully benefit from the vertex detector the tracking data had to be used to improve the single hit resolution to 5 μm or less, close to the true intrinsic resolution.

A major goal of the project was to find a prescription for combining the alignment data from particle tracks in a deterministic way, so that a robust optimum geometry could be obtained, i.e., one which does not depend on the vagaries of the order in which parameters were optimized, nor require many subjective judgements to be made. We believe that this goal has been achieved and will describe the mathematical underpinnings of the result.

A key tool which has enabled us to achieve the result was the use of singular value decomposition (SVD). We explain how the SVD technique was integrated into our alignment procedure to find an optimum χ^2 minimization solution for the CCD geometry using the track data. In the following section we give an overview of the method. In Section 7.3 we describe how the displacements of a CCD from its nominal location perturbs the position at which a track hit is observed. Section 7.4 discusses how hits from various types of track data are included and the constraints made on the detector geometry. A description is then given in Section 7.5 of how these constraints are combined into a global set of equations which have been solved using SVD. In Section 7.6 examples are given of the difference in the distributions of track residuals before and after the alignment. Finally in Sections 7.7 and 7.8 we make some concluding remarks and discuss how the technique could be applied in the alignment of other trackers.

A fuller description of the method and the development of the alignment, plus a brief overview of SVD, can be found in Refs. [5, 6].

7.2 Overview of the method

Although the CDC was used to identify charged tracks, most of the observables used in the internal alignment were calculated using just CCD hits. Various residuals were defined, each of which can be described as a simple (linear) function of the misalignments of the CCDs used in that particular residual, and a number of parameters which can be determined with sufficient accuracy from the nominal geometry of the detector and approximate track parameters.

For the initial residuals studied, good quality tracks were constrained to pass through two of the CCD hits and the residual measured to the third, reference, CCD. Later more complex residuals were included, each using up to four CCDs, as described below.

For each type of residual and given combination of CCDs a set of tracks to be used in the alignment were assembled. The measured residuals for each set can in general be expressed as a polynomial of known track parameters (here values from the nominal geometry are adequate), and the unknown CCD misalignments enter only in the coefficients of the various polynomial terms. Thus fitting the measured residuals to the appropriate polynomial allows each set of tracks to be reduced to a few coefficients. It is then possible to write a matrix equation **Ax = c**, where **x** is a vector of the n unknown CCD misalignments, **c** is a vector of the m coefficients fitted to the various sets of tracks and **A** is an $m \times n$ matrix with elements determined from the nominal geometry. This equation can be solved to give the required CCD misalignments, however, this does not take into

account the errors and correlations in the residual fits. A better solution can be obtained by redefining the basis of coefficients from each of the residual fits such that the modified coefficients from each fit have a unit covariance matrix.

The equation to be solved then becomes **TAx = Tc**, where **T** provides the suitable change of basis for the coefficients **c**. The elements of **T** are calculated from the covariance matrices of the residual fits. Solving this equation with SVD effectively gives a more optimal χ^2 fit for the unknown CCD misalignments **x**.

7.3 CCD misalignments and hit residuals

Figure 7.2 shows the degrees of freedom allowed in this analysis for a pair of CCDs on one of the 48 ladders of VXD3.

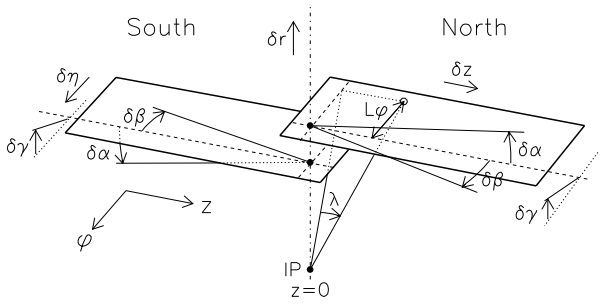

Fig. 7.2: Definition of the CCD translation δz, δr, $\delta\eta$ and rotation $\delta\alpha$, $\delta\beta$, $\delta\gamma$ corrections to be determined, indicated for two CCDs on the same ladder

The parameters consist of three translations, parallel to the edges of the rectangular CCD and normal to the CCD plane, and three rotations about the normal point in the CCD plane (i.e., the point in the plane closest to the IP)

- δz : Translation in CCD plane in z direction
- $\delta\eta$: Translation in CCD plane in ϕ direction
- δr : Translation normal to CCD
- $\delta\alpha$: Pitch – Rotation axis along width of CCD
- $\delta\beta$: Yaw – Rotation axis normal to CCD
- $\delta\gamma$: Roll – Rotation axis along length of CCD.

Figure 7.2 also shows how the hit • in a CCD is specified by the angle λ in the rz plane relative to $z = 0$ about the IP, and the distance L_ϕ across the CCD in the CCDs own reference system. The alignment procedure described here assumed that each CCD was approximately flat, with small shape corrections as measured in the optical survey having been applied (see further discussion on CCD shapes in Section 7.6).

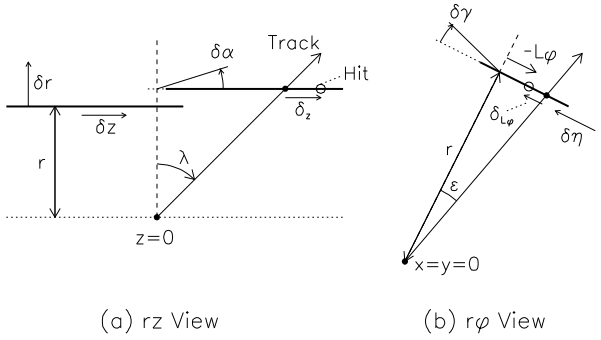

Fig. 7.3: The effect on the apparent hit position in a CCD due to adjustments in the six degrees of freedom

Figure 7.3 illustrates the effect on the apparent hit position within a CCD for a track of fixed trajectory when the CCD position is adjusted by movements in the six degrees of freedom. Misalignments of the CCDs cause the measured hit on the CCD to be displaced from the true track trajectory by a residual amount δ_z along the CCD length and $\delta_{L\phi}$ across its length. The sign of δ_z is such that it measures the z location of the hit minus the z location of the actual track in the plane of the CCD. If the only degree of freedom of the CCD was the δz correction, then $\delta_z = -\delta$z, and it would be trivial to 'unfold' the required CCD alignment correction δz from the measured residual δ_z. Straightforward geometric arguments show that the more general form for the δ_z and $\delta_{L\phi}$ residuals can be approximated as:

$$\delta_z = -\delta\mathrm{z} + \delta\mathrm{r}\tan\lambda + \delta\alpha\,\mathrm{r}\tan^2\lambda + \delta\gamma L_\phi\tan\lambda + \delta\beta L_\phi \tag{7.1}$$

$$\delta_{L\phi} = -\delta\eta + \frac{\delta\mathrm{r}}{\mathrm{r}}L_\phi + \frac{\delta\gamma}{\mathrm{r}}{L_\phi}^2 + \delta\alpha L_\phi\tan\lambda - \delta\beta\mathrm{r}\tan\lambda \tag{7.2}$$

where both residuals are measured in the plane of the CCD.

Since the true track trajectory is unknown it is necessary to identify hits on several (usually three) CCDs associated with a track reconstructed in the CDC. Good quality tracks were selected with a momentum of at least 1 GeV. In general the track was constrained to pass through two of the CCD hits and the corresponding residual measured to the third, reference, CCD. The small curvature effect of the charged track of known momentum in the SLD magnetic field (0.6 T) was taken into account in the $r\phi$ plane. The relative lever-arm weights with which each of the three CCDs contribute to the observed residual was determined from the ideal geometry to within a very good approximation (since the alignment corrections are very small compared with the dimensions of the detector), in all cases the refer-

ence CCD was given a weight $= +1$. For example, an observed δ_z residual is equated with the sum of three weighted CCD contributions of the form of Eq. (7.1). The CCD contributions to the $\tan\lambda$ terms in Eqs. (7.1) and (7.2) are readily combined since $\tan\lambda$ is a parameter of the track itself which is common to all hits. Terms involving L_ϕ require more care since in general this value differs for each CCD even for the same track. However, we determine L_ϕ on the reference CCD and to a close approximation the value of this local coordinate on another CCD i, L_{ϕ_i}, is linearly related as $L_{\phi_i} = a_i L_\phi + b_i$ where a_i and b_i are constant factors determined solely by the ideal geometry of the CCDs concerned. Taking all of these terms into account leads to the following general functional forms for the residual measured on the reference CCD for a given track:

$$\begin{aligned}\delta_z = \sum_i w_i[\quad & (-\delta \mathrm{z}_i + b_i\delta\beta_i) && \mathbf{1} \\ + \quad & (\delta \mathrm{r}_i + b_i\delta\gamma_i) && \boldsymbol{\tan\lambda} \\ + \quad & (\mathrm{r}_i\delta\alpha_i) && \boldsymbol{\tan^2\lambda} \\ + \quad & (a_i\delta\gamma_i) && \boldsymbol{L_\phi \tan\lambda} \\ + \quad & (a_i\delta\beta_i) && \boldsymbol{L_\phi}]\end{aligned} \tag{7.3}$$

$$\begin{aligned}\delta_{L\phi} = \sum_i w_i f_i[\quad & (-\delta\eta_i + \tfrac{b_i}{\mathrm{r}_i}\delta \mathrm{r}_i + \tfrac{b_i^2}{\mathrm{r}_i}\delta\gamma_i) && \mathbf{1} \\ + \quad & (\tfrac{a_i}{\mathrm{r}_i}\delta \mathrm{r}_i + \tfrac{2a_ib_i}{\mathrm{r}_i}\delta\gamma_i) && \boldsymbol{L_\phi} \\ + \quad & (\tfrac{a_i^2}{\mathrm{r}_i}\delta\gamma_i) && \boldsymbol{{L_\phi}^2} \\ + \quad & (a_i\delta\alpha_i) && \boldsymbol{L_\phi \tan\lambda} \\ + \quad & (b_i\delta\alpha_i - \mathrm{r}_i\delta\beta_i) && \boldsymbol{\tan\lambda}]\end{aligned} \tag{7.4}$$

where the sum is over the contributing CCDs and their signed weights w_i were determined relative to the reference CCD, according to the geometric lever arm. For example, the layer 2 reference CCD of a Triplet (described in Section 7.4) is given a weight of $+1.0$ while the layer 1 and layer 2 CCDs each contribute with a weight $w_i = -0.5$. Corrections to the non-reference CCD contributions due to the fact that the CCDs are not parallel in the $r\phi$ plane are determined from the ideal geometry and yield the factors $f_i(\sim 1.0)$. For most of the residual types a reduced functional form can be used and these are shown in Table 7.1 and described in the following section.

7.4 Residual types, distributions and fits

The internal alignment began by classifying the types of tracking constraints as shown in Fig. 7.4 with the corresponding residuals δ indicated in the rϕ plane. In each case there are analogous residuals in the rz plane. There are 48 pairs of CCDs shown in Fig. 7.4 with the inner and outer lines representing the North and South CCDs, respectively. The six residual types used are described below.

- Shingles - require a pair of hits in adjacent CCDs within the same layer, plus a third, anchor, hit in another layer. These constrain the relative positions of CCDs within a layer.
- Doublets - require a pair of hits in the two CCDs on a single ladder (in the overlap region halfway along the length), plus a third hit in another layer. These constrain the relative positions of the two halves of VXD3.
- Triplets - require three hits in different layers, either all on the North or all on the South half of the detector. These constrain the relative positions of the three layers.
- Pairs - require a pair of back-to-back tracks from $Z^0 \to \mu^+\mu^-$ or $Z^0 \to e^+e^-$ events, each giving a hit in layers 1 and 3 (layer 2 is ignored for Pairs). These constrain opposite regions of the detector.
- VXD3 vs. CDC Track Angle Matching - the angle of high momentum tracks measured in the CDC is compared with the angle measured in layers 1 and 3 of VXD3. The angular discrepancy is converted into a residual distance on a reference CCD. The residual uncertainty was determined from the CDC track as well as the CCD hit resolution.
- IP constraint - on average tracks measured in layers 1 and 3 should project back through the real interaction point. For this constraint the IP is allowed to float in the xy plane. In addition to the 6×96 CCD corrections a further two, δx and δy, are introduced for the deviation of the IP from the nominal position. Hence there are a total of 578 alignment corrections to be determined. The contribution to the constant term i_1^+ near the bottom of Table 7.1 is given by:

$$i_1^+ = f(\delta_{\mathrm{CCD}}) - \delta x \sin\phi_i + \delta y \cos\phi_i \tag{7.5}$$

where $f(\delta_{\mathrm{CCD}})$ is determined using Eq. (7.4) and ϕ_i is a constant taken to be the average value of ϕ for each residual region.

The full set of functional forms is listed in Table 7.1.

Table 7.1: Functional forms for fitting the various residual types, N_I lists the total number of independent residual fits involved while N_C is the number of coefficients determined

Type	Functional form	N_I	N_C
Shingles	$\delta_z^s = s_1^{\parallel} + s_2^{\parallel} \tan\lambda + s_3^{\parallel} \tan^2\lambda$	96	288
	$\delta_{L\phi}^s = s_1^{\perp} + s_2^{\perp} \tan\lambda$	96	192
Doublets	$\delta_z^d = d_1^{\parallel} + d_2^{\parallel} L_\phi$	48	96
	$\delta_{L\phi}^d = d_1^{\perp} + d_2^{\perp} L_\phi + d_3^{\perp} {L_\phi}^2$	48	144
Triplets	$\delta_z^t = t_1^{\parallel} + t_2^{\parallel} \tan\lambda + t_3^{\parallel} \tan^2\lambda + t_4^{\parallel} L_\phi \tan\lambda + t_5^{\parallel} L_\phi$	80	400
	$\delta_{L\phi}^t = t_1^{\perp} + t_2^{\perp} L_\phi + t_3^{\perp} {L_\phi}^2 + t_4^{\perp} L_\phi \tan\lambda + t_5^{\perp} \tan\lambda$	80	400
Pairs	$\delta_{rz}^p = p_1^{\parallel} + p_2^{\parallel} \tan\lambda + p_3^{\parallel} \tan^2\lambda$	28	84
	$\delta_{r\phi}^p = p_1^{\perp} + p_2^{\perp} \tan\lambda$	28	56
	$\delta_{\phi}^p = p_1^{\phi} + p_2^{\phi} \tan\lambda$	28	56
CDC angle	$\delta_{\lambda}^c = c_1^{\lambda} + c_2^{\lambda} \tan\lambda + c_3^{\lambda} \tan^2\lambda$	56	168
	$\delta_{\phi}^c = c_1^{\phi} + c_2^{\phi} \tan\lambda$	56	112
IP constraint	$\delta_{IP}^i = i_1^{\perp} + i_2^{\perp} \tan\lambda$	56	112
	Total	700	2108

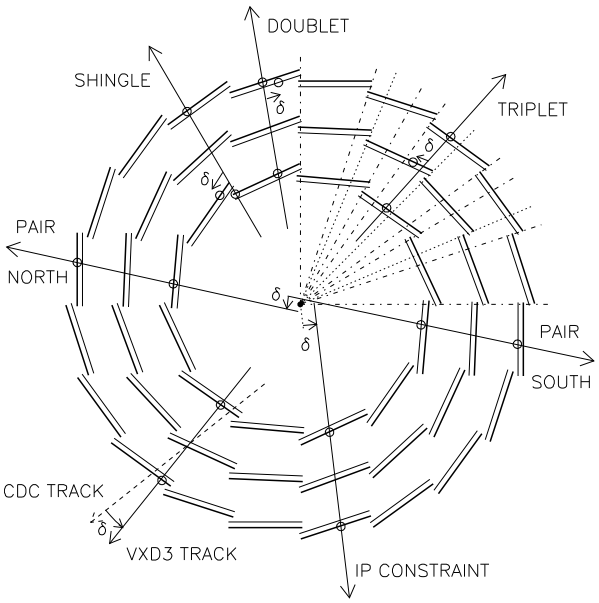

Fig. 7.4: Construction of the VXD3 residuals

The dashed and dash-dotted lines in the upper right quadrant of Fig. 7.4 indicate the division of the data into ten Triplet regions, while the dash-dotted lines only indicate the division into seven regions used to define the Pair, CDC angle matching and IP constraint residual regions. These boundaries are determined by the requirement that each residual region involves a unique set of CCDs. These divisions repeat in the other three quadrants by the symmetry of the detector. The 40 Triplet regions correspond to the $N_I = 80$ independent fits listed in Table 7.1 since North and South CCDs are considered separately for the Triplets.

For each type of residual and each unique combination of CCDs n-tuples were accumulated containing the deviations δ_z and $\delta_{L\phi}$, and co-ordinates $\tan\lambda$ and L_ϕ. These n-tuples were fitted to the functional forms given in Table 7.1 to determine the coefficients of the deviations and the covariance error matrix of the fit. These fits were done using MINUIT [7] with an automated procedure to loop over the large number of residual type and CCD combinations involved. Examples of the fits for a Shingle are shown in Fig. 7.5.

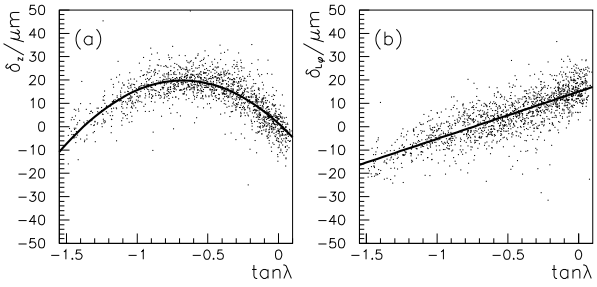

Fig. 7.5: An example of (a) δ_z and (b) $\delta_{L\phi}$ residual distributions as a function of $\tan\lambda$ for a Shingle in layer 2 before alignment. The solid lines show the fitted curves.

Each of the points in Fig. 7.5 represents the residual obtained from one track. All residual fits were made to data points obtained by binning the distributions in $\tan\lambda$ (and/or L_ϕ) and averaging over the track data in each bin to determine the data point with an error bar, taking the intrinsic hit resolution into account. After each fit the automated procedure included the search

for and removal of outliers (assumed to be caused by mis-linked tracks) and a refit to the remaining data if necessary. This was particularly important for fits with relatively little track data, and in all cases a visual cross-check was made using a graphical fit display. The scatter of the points about the fitted curves in Fig. 7.5 is due to the intrinsic resolution.

The effects of the misalignments to be corrected are seen as a parabola and non-zero gradient slope in Fig. 7.5(a) and (b) respectively, as predicted by the corresponding functional forms for Shingles listed at the top of Table 7.1.

7.5 The alignment matrix equation and solution with SVD

The repeated application of Eqs. (7.3) and (7.4) yields a large set of simultaneous equations which can be expressed in terms of a single matrix equation $\mathbf{Ax} = \mathbf{c}$, where $\mathbf{x}$ is a column matrix of the 578 CCD and IP degrees of freedom, $\mathbf{c}$ is a column matrix of coefficients measured in the residual fits listed in Table 7.1 and the 2108×578 elements of $\mathbf{A}$, called the 'design matrix' are determined by the ideal geometry:

$$\begin{pmatrix} \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \ldots\mathbf{A}(w_i, r_i, a_i, b_i, f_i)\ldots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} \begin{pmatrix} \delta \mathrm{z}_1 \\ \vdots \\ \delta \mathrm{z}_{96} \\ \delta\eta_j \\ \delta \mathrm{r}_j \\ \delta\alpha_j \\ \delta\beta_j \\ \delta\gamma_j \\ \delta x \\ \delta y \end{pmatrix} = \begin{pmatrix} s_1^{\parallel}(96) \\ s_2^{\parallel}(96) \\ s_3^{\parallel}(96) \\ s_1^{\perp}(96) \\ s_2^{\perp}(96) \\ d_1^{\parallel}(48) \\ \vdots \\ t_1^{\parallel}(80) \\ \vdots \\ \vdots \\ p_1^{\parallel}(28) \\ \vdots \\ c_1^{\lambda}(56) \\ \vdots \\ i_1^{\perp}(56) \\ i_2^{\perp}(56) \end{pmatrix} \tag{7.6}$$

where, for example, the $\delta\eta_j$ denote the 96 CCD corrections $\delta\eta$ to be determined. With $\mathbf{A}$ and $\mathbf{c}$ known the question is how to find the optimal solution to this equation for the alignment corrections $\mathbf{x}$. Several features of matrix singular value decomposition are listed in the appendix of Ref. [5]. However, the important property of interest here is that by inverting the possibly singular $m \times n$ matrix $\mathbf{A}$ using SVD yields the $m \times n$ matrix $\mathbf{A}^{+}$, such that the solution $\mathbf{x} = \mathbf{A}^{+}\mathbf{c}$ minimizes the magnitude $\mid \mathbf{Ax} - \mathbf{c} \mid$. This 'least squares fit' minimization of $\sum_i (\sum_j A_{ij}x_j - c_i)^2$ could directly be used to determine an alignment geometry, however, it does not take into account the error and correlation information contained in the covariance matrices of the residual fits. It is necessary to redefine the basis of parameters for each residual fit such that each covariance matrix takes a unit diagonal form for the least squares fit using SVD to effectively become a more optimal χ^2 fit.

Quite generally the $n \times n$ covariance error matrix $\mathbf{V}$ for a fit to n parameters of vector $\mathbf{p}$ can be decomposed as $\mathbf{V} = \mathbf{HH}^T$ where $\mathbf{H}$ is a non-singular lower triangular matrix. The new vector $\mathbf{p}' = \mathbf{H}^{-1}\mathbf{p}$ then contains a set of n independent Gaussian variables with a unit covariance error matrix [8]. We use the CERNLIB [9] routines TRCHLU to decompose the error matrix of each residual fit $\mathbf{V} = \mathbf{HH}^T$ and TRINV to invert the triangular matrix $\mathbf{H} \rightarrow \mathbf{H}^{-1}$. Multiplying each $\mathbf{H}^{-1}$ matrix into the relevant elements of $\mathbf{c}$ removes the residual fit correlations and with each of the new elements having an error of unity. Symbolically we collect all of the non-zero elements of the 700 triangular $\mathbf{H}^{-1}$ matrices into a single 2108×2108 matrix denoted $\mathbf{T}$ and in order to maintain the equality in Eq. (7.6) multiply into both sides:

$$\mathbf{TAx} = \mathbf{Tc} \tag{7.7}$$

yielding:

$$\mathbf{A}'\mathbf{x} = \mathbf{c}', \tag{7.8}$$

where the rows of $\mathbf{A}'$ and elements of $\mathbf{c}'$ have been obtained by the same linear transformation, under the operation of $\mathbf{T}$, from the rows of $\mathbf{A}$ and elements of $\mathbf{c}$, respectively. The new $\mathbf{A}'$ matrix is then decomposed according to the SVD prescription and the inverse $\mathbf{A}'^{+}$ constructed. Finally, the solution $\mathbf{x} = \mathbf{A}'^{+}\mathbf{c}'$ determines a χ^2 minimization over all the track residual data and yields the 578 alignment corrections.

7.6 Results of vertex detector alignment

7.6.1 Including CCD shape corrections

After following the prescription discussed in the previous sections and correcting the geometry according to the SVD solution, optimal resolution in the $r\phi$ plane was effectively achieved. However the rz performance was less satisfactory, particularly for high $|\tan\lambda|$ where tracks make a large angle to the CCD normal and the resolution becomes more susceptible to the radial uncertainty in the CCD hit location. From the optical survey data the shapes of the CCD surfaces had been fitted to 14 parameter Chebychev polynomials [2]. In particular, the CCDs were observed to take a 4^{th} order polynomial 'W' shape, seated on their glue attachments to the ladders, along the z-direction, with deviations from a flat surface typically a few tens of microns in size (see Fig. 7.6). This systematic distortion of mechanical origin was expected to be modified under the detector cooldown from room to operating temperature. In fitting for the six alignment corrections for each of the 96 CCDs it had been assumed that the shapes were perfectly known from the optical survey. Examination of

the structure of individual residual region distributions showed deviations from the assumed functional forms listed in Table 7.1, consistent with the 'W' shapes not having been fully corrected due to both the limited survey data resolution and the thermal difference between the room-temperature survey and the VXD3 cryostat. The effects seen in the residual distributions, indicated that the resolution might be further improved by applying shape corrections based on the track data.

With the δz and $\delta\alpha$ rigid body CCD corrections already accounting for two of the five variables required to parameterize a 4^{th} order polynomial, three further shape-correcting terms δq, δh and δt (representing the displacement from a flat surface at $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$ of the distance along a CCD with both ends fixed) were introduced.

The shape was treated as a $\tan\lambda$ dependent radial correction: $\delta\mathrm{r} \rightarrow \delta\mathrm{r} + f(\delta\mathrm{q}, \delta\mathrm{h}, \delta\mathrm{t}, \tan\lambda)$, where the function f was determined by a few lines of straightforward algebra.

Substituting this redefined expression for δr into Eqs. (7.1) and (7.2) and following the complete procedure described in the previous three sections with the addition of the shape-correcting terms led to a new version of the matrix equation $\mathbf{Ax} = \mathbf{c}$ of Eq. (7.6). This consisted of a 4160×866 design matrix $\mathbf{A}$ determined from the ideal geometry, an array of 866 unknown alignment corrections $\mathbf{x}$ (now including the δq, δh and δt displacements for each CCD) and an array of 4160 coefficients $\mathbf{c}$ determined from the original 700 residual fits with new higher order terms accounting for the CCD shapes.

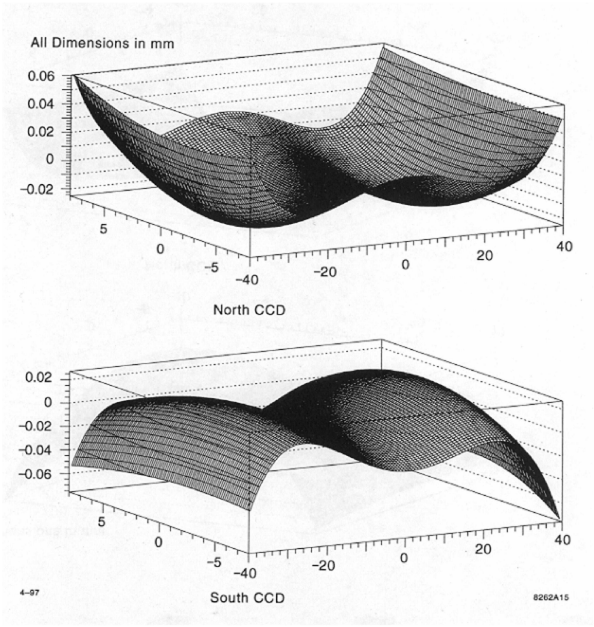

Fig. 7.6: Example of typical shapes of the two CCDs on a ladder, as measured in the room-temperature optical surveys

Table 7.2: Functional forms for fitting the various residual types after the shape corrections were included, N_I lists the total number of independent residual fits involved while N_C is the number of coefficients determined

Type		**Terms in functional forms**													N_I	N_C
		1	$\tan\lambda$	$\tan^2\lambda$	$\tan^3\lambda$	$\tan^4\lambda$	$\tan^5\lambda$	L_ϕ	$L_\phi{}^2$	$L_\phi\tan\lambda$	$L_\phi\tan^2\lambda$	$L_\phi\tan^3\lambda$	$L_\phi\tan^4\lambda$			
Shingles	$\delta^s_z =$	$s^\parallel_1$	$s^\parallel_2$	$s^\parallel_3$	$s^\parallel_4$	$s^\parallel_5$	$s^\parallel_6$								96	576
	$\delta^s_{L\phi} =$	$s^\perp_1$	$s^\perp_2$	$s^\perp_3$	$s^\perp_4$	$s^\perp_5$									96	480
Doublets	$\delta^d_z =$	$d^\parallel_1$						$d^\parallel_2$							48	96
	$\delta^d_{L\phi} =$	$d^\perp_1$						$d^\perp_2$	$d^\perp_3$						48	144
Triplets	$\delta^t_z =$	$t^\parallel_1$	$t^\parallel_2$	$t^\parallel_6$	$t^\parallel_7$	$t^\parallel_8$		$t^\parallel_5$	$t^\parallel_4$						80	640
	$\delta^t_{L\phi} =$	$t^\perp_1$	$t^\perp_5$	$t^\perp_9$	$t^\perp_{10}$	$t^\perp_{11}$		$t^\perp_2$	$t^\perp_3$	$t^\perp_4$	$t^\perp_6$	$t^\perp_7$	$t^\perp_8$		80	880
Pairs	$\delta^p_{rz} =$	$p^\parallel_1$	$p^\parallel_2$	$p^\parallel_3$	$p^\parallel_4$	$p^\parallel_5$	$p^\parallel_6$								28	168
	$\delta^p_{r\phi} =$	$p^\perp_1$	$p^\perp_2$	$p^\perp_3$	$p^\perp_4$	$p^\perp_5$									28	140
	$\delta^p_\phi =$	p^ϕ_1	p^ϕ_2	p^ϕ_3	p^ϕ_4	p^ϕ_5									28	140
CDC angle	$\delta^c_\lambda =$	c^λ_1	c^λ_2	c^λ_3	c^λ_4	c^λ_5	c^λ_6								56	336
	$\delta^c_\phi =$	c^ϕ_1	c^ϕ_2	c^ϕ_3	c^ϕ_4	c^ϕ_5									56	280
IP constraint	$\delta^i_{IP} =$	$i^\perp_1$	$i^\perp_2$	$i^\perp_3$	$i^\perp_4$	$i^\perp_5$									56	280
														Total	700	4160

The functional forms used after the shape corrections had been included are listed in Table 7.2. Since there was no track data to constrain the shape in the high $|\tan\lambda|$ region of layer 2 and particularly layer 1 CCDs (as can be seen from Fig. 7.1(a) tracks from the IP that traverse any part of a layer 3 CCD can only intersect about half of a layer 1 CCD), a further 866 rows, each with only one non-zero element, were appended to the matrix **A** and the array **c** was extended downwards with 866 zero elements. These extra lines were used to apply conservative restraints on the geometry by effectively combining the dummy measurements δz, δr $= 0 \pm 10$ μm, $\delta\eta = 0 \pm 2$ μm, $\delta\alpha, \delta\beta, \delta\gamma = 0 \pm 0.05$ mrad and $\delta q, \delta h, \delta t = 0 \pm 5$ μm (δx and δy for the IP were not ultimately constrained in this way) into the SVD χ^2 minimization fit. The complete 5026×866 element matrix $\mathbf{A}'$ and the 5026 element vector $\mathbf{c}'$ were obtained from **A** and **c**, respectively, by Eq. (7.7) using the 16 332 non-zero elements of the effective matrix **T**. The latter was obtained in turn from the residual fit covariance error matrices (plus the corresponding error terms from the dummy measurements added above) as described in the previous section. The matrix $\mathbf{A}'$ was inverted using SVD and the VXD3 geometry was corrected with the 866 elements of the solution vector $\mathbf{x} = \mathbf{A}'^{+}\mathbf{c}'$.

7.6.2 Achievement of design performance

The alignment procedure described in the previous sections requires only a single iteration to a given data set to determine the corrected geometry. In practice, due to a difference in run conditions and as the algorithm developed, several aligned geometries were determined for the data taken in 1996/98. Details of these can be found in Ref. [6].

Figure 7.7 shows the Triplet δ_z and $\delta_{L\phi}$ residuals obtained with the aligned detector together with the pre-alignment residuals derived from the optical survey geometry. Since the Triplets cover the full volume of the detector, that is over the full region of track acceptance for all CCDs in all three layers, they provide a representative indication of the local alignment. The data used for these Triplet plots correspond to charged tracks selected with momentum greater than 5 GeV/c to suppress the multiple scattering contribution.

The post-alignment RMS of the residual distributions was found to be around a factor of four improved over the pre-alignment RMS values. Similar plots for all residual types, along with the mean and RMS measurements, were a major guide in debugging and refining the alignment algorithm. Indeed, only after taking into account the full residual fit error matrices, with the operation of **T** as expressed in Eq. (7.7), were the post-alignment residual plots observed to stand out in a manner approaching the ideal performance, i.e., with an RMS dominated by the intrinsic CCD hit resolution about an essentially zero mean as seen in Fig. 7.7. Fitting a single Gaussian curve to each of the post-alignment histograms in Figs 7.7(a) and (b) yielded a width of 4.45 μm in each case. This number, depending on the three Triplet CCDs, is divided by the geometric weight factor $\sqrt{1.0^2 + 0.5^2 + 0.5^2}$ to yield 3.63 μm as the single CCD hit resolution for both δ_z and $\delta_{L\phi}$ after the alignment, well within the initial target goal of 5 μm and close to the intrinsic hit resolution. Erratic variations in the residual distribution mean as a function of the Triplet index that had been observed in the initial data were reduced to the 1 μm level, as can be seen in Fig. 7.8.

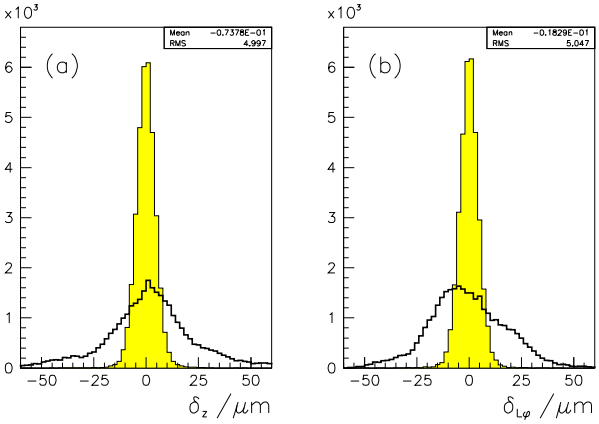

Fig. 7.7: Triplet residuals for 1997/1998 data obtained with the original survey geometry (broad histograms with thick outline) and after the alignment (narrow shaded histograms)

The magnitudes of the measured geometry corrections were typically O(10 μm) with the largest effects being perpendicular to the CCD plane. A second iteration of the global alignment, described in Section 7.1, was performed after the internal alignment and gave only minor corrections for the six rigid body degrees of freedom of the detector as a whole. It should be noted that the SVD procedure does not need to be iterated on the same data set. This was expected given the relatively minor nature of the approximations made in the analysis and was confirmed by the observed negligible effect of performing a second iteration of the internal alignment.

After the final alignment, the one hit resolution for the Shingle, Doublet, Triplet and Pair residuals over the whole detector was in general found to be consistently better than 4.0 μm in both the rz and $r\phi$ planes, very close to the true intrinsic CCD resolution, and design performance had been achieved.

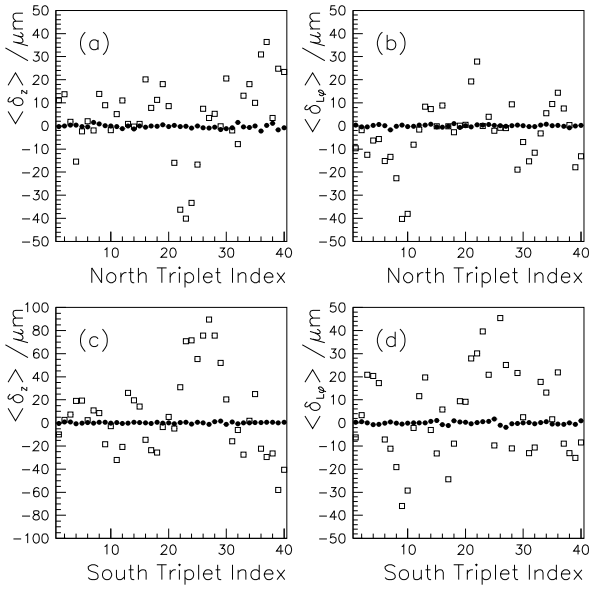

Fig. 7.8: Triplet δ_z and $\delta_{L\phi}$ residual mean as a function of the ϕ-dependent Triplet index (see Fig. 7.4), for the North and South halves of VXD3, obtained before $\Box$ and after $\bullet$ the alignment corrections (plot (c) made with different vertical scale).

7.7 Conclusions on the alignment

With the internal alignment of the vertex detector complete the charged track parameters were determined using a combined fit to the VXD3 hits and CDC (Central Drift Chamber) track data. Compared with the VXD3 only Pair data discussed above, the impact parameter resolution improves due to the inclusion of the layer 2 VXD3 hit and further in the $r\phi$ plane due to the CDC information. The constant and momentum/angle dependent multiple scattering terms for the final rz and $r\phi$ impact parameter resolution at the IP in the 1997/1998 SLD data were determined to be:

$$\begin{aligned} \sigma_{rz} &= 9.7 \oplus \frac{33}{p \sin^{3/2}\theta}\ \mu\text{m} \\ \sigma_{r\phi} &= 7.8 \oplus \frac{33}{p \sin^{3/2}\theta}\ \mu\text{m} \end{aligned} \qquad (7.9)$$

where the constant terms depend primarily upon the VXD3 CCD intrinsic resolution and geometry. In order to achieve this design performance a great deal of effort outside of the SVD analysis presented here was required. This work included the initial optical survey and corrections for electronics and mechanical effects only briefly mentioned in this paper (for more details see Ref. [2]).

Ultimately, however, only the SVD analysis had the necessary versatility to simultaneously correct the geometry of each of the 96 individual CCDs for three translational, three rotational, and three shape-defining degrees of freedom. The variety of residual types, including the Pair, CDC matching angle and IP constraint data, successfully incorporated into the SVD fit further demonstrate the remarkable robustness of the technique. The quality of the solution depended on the careful inclusion of the residual fit error information into the analysis, as described in Section 7.5 and Eq. (7.7), to adapt the least squares fit property of the SVD method into an optimal χ^2 minimization fit. The algorithm is extremely robust with little need for 'tuning' except in the treatment of effects where the real data is observed to deviate significantly from the ideal model and any assumptions used in defining the problem.

Although the size of the matrices involved appears daunting, only $\sim 1\%$ or $\sim 35\,000$ elements of the 5026×866 matrix **A** were given non-zero values, and these in turn were generally determined in a straightforward way from only ~ 10 parameters describing the VXD3 geometry in Fig. 7.1. Also, the large number of non-zero elements involved in the **T** matrix and **c** array were obtained from the residual data skims by automated looping over large numbers of similar fits each identified by an index label. Much of the analysis was an exercise in book-keeping.

While only one iteration of the whole analysis was required to obtain an optimal solution, we found some benefit from tuning the geometry restraints discussed at the end of the last section. Applying a trivial dummy measurement ($\delta z = 0 \pm 10\mu$m, etc.) to each alignment correction gives a fully constrained set of alignment equations. This was of value early in the analysis, for example, using the Shingle residual fit data only together with the dummy restraints allowed a clean test of the code relevant for the Shingle data only without interference from the other data types. We also found some benefit in artificially reducing the residual fit errors by tuning an extra weight in matrix **A** for residual types that suffered from low statistics, for example, the Pair data in the early run period. Such fine-tuning, applied toward the end of the analysis, did not play a significant role.

A further iteration of the whole procedure was only necessary after a significant change in the data sample or an improvement in the algorithm such as the inclusion of the IP constraint or CCD shape-correcting terms. The project began with the Monte Carlo studies as the first VXD3 data was being collected in 1996, was refined as the data accumulated, and concluded in May 1999 within a year of the end of the final SLD run. Although some visible effects in the residual data remained, for example, due to a slight bow shape across the CCD width observed in a few Doublets, these were

all too minor to consider further extensions of the algorithm. The achievement of the design resolution indicated in Eq. (7.9) has greatly benefitted the rich programme of heavy flavour physics analysis at SLD during the subsequent years.

7.8 Lessons for other trackers

The basic technique used could be applied for other systems where: the required solution is a perturbation described by O(1000) parameters which are small compared to the dimensions of the system; constraining data exists which can be expressed as a set of simultaneous linear equations of the parameters. Clearly, as for any method, the constraining data has to give adequate coverage, not only for local corrections, but also for longer distance corrections (sometimes referred to as 'weak modes').

This work has demonstrated that it was already possible in 1999 to handle simply and reliably the matrices required in the case of VXD3 (inversion of sparse matrices of order 5000×1000 elements) using double precision arithmetic in modest times on a standard workstation. In this case only $\sim 1\%$ of the elements of the design matrix had non-zero values. Given improvements since then in computer hardware and matrix manipulation code it should now be possible to handle problems several times larger.

Using Singular Value Decomposition provided a robust technique for the matrix inversion and also provided a measure of the degrees of freedom left unconstrained by the constraint data. However, the basic technique could be modified to use other matrix inversion techniques, to give similar statistical dependence on the data and the geometry.

The alignment problem of a tracker is greatly assisted by a suitable design of the detector itself. Here we would list as significant properties:

- Symmetry of the detector - this greatly assists book-keeping and allows comparison of different parts of the detector.
- Overlap regions - allow devices to be stitched together with a favourable lever arm.
- Large devices - obviously it is better to have a single element (with fewer degrees of freedom) than two with an overlap. Provided of course that the single element has a stable shape with corrections that can be described with a few parameters. Indeed we would argue that it may be preferable to use detectors where a deliberate bow has been introduced if this leads to greater mechanical stability and decreased shape uncertainty compared to using flat detectors.
- Stability - the geometry (both devices and support structure) should be stable with respect to time. Thus changes due to temperature fluctuations, cycling of magnetic fields, ageing under gravity/elastic forces, etc., should be 'small', at least over the period of time long enough to collect sufficient track data for the alignment.

Before the technique could be successfully applied to a different tracker it would be necessary to identify the required constraint data, with sufficient coverage to resolve all of the important degrees of freedom. This could, as in our case be an iterative process, adding additional terms and constraints as indicated by deviations in the results. Indeed it is vital that the various residuals used are replotted after applying the results of the alignment, and checked for any systematic deviations. Such systematic deviations can give vital clues to errors in the models used and extra terms which might be needed. In addition, if the results are being used in the calculation of derived quantities (such as masses) these should also be checked for systematic problems using known references, again indicating the need for additional constraints. It is important that the constraint data includes not only data to fix local affects (i.e., relative alignment of neighbouring elements), but also data to fix intermediate and long-range affects. In our case we used back-to-back pairs, CDC matching angle and IP constraints, but cosmic ray tracks could be another rich source of such constraint data.

Acknowledgements

While this paper has described our use of the singular value decomposition matrix technique, the successful VXD3 alignment and performance has been the responsibility of a large number of SLD Collaborators [2]. In particular, we would like to acknowledge the MIT group for performing the initial optical survey, John Jaros for coordinating the VXD3 tracking program, Danning Dong for work on the Monte Carlo study and global alignment, Aaron Chou for making offline reconstruction work and for performing the IP position finding, Nikolai Sinev for correcting electronics hardware problems before the 1997 run, and Chris Damerell for discussions on the mechanical and thermal properties of the VXD3 components.

This work was supported by the UK Particle Physics and Astronomy Research Council, and the US Department of Energy contract DE-AC03-76SF00515. The greater part of this paper is derived from an earlier description of the work published in Ref. [5]. Some details have been omitted in this paper, whilst some clarifications and additional material have been included.

References

[1] D.J. Jackson, *Nucl. Instrum. Methods* **A388** (1997) 247.

[2] K. Abe *et al.*, *Nucl. Instrum. Methods* **A400** (1997) 287.

[3] The CCDs were manufactured by the EEV Company, Chelmsford, Essex, United Kingdom.

[4] M.J. Fero *et al.*, *Nucl. Instrum. Methods* **A367** (1995) 111.

[5] D.J. Jackson, D. Su and F.J. Wickens, *Nucl. Instrum. Methods* **A510** (2003) 233.

[6] D.J. Jackson, D. Su and F.J. Wickens, SLD Note 271.

[7] F. James and M. Roos, *Comput. Phys. Commu.* **10** (1975) 343.

[8] Particle Data Group, *Phys. Rev. D* **50** (1994) 1273.

[9] *CERNLIB - The CERN computer centre program library*, `http://cernlib.web.cern.ch/cernlib/` .