

Global χ^2 approach to the Alignment of the ATLAS Silicon Tracking Detectors

P. Brückman

University of Oxford
DWB, Keble Road
Oxford OX1 3RH, UK
and

Institute of Nuclear Physics PAN
ul. Radzikowskiego 152 Kraków 31-342, Poland
e-mail: P.Bruckman1@physics.ox.ac.uk

A.Hicheur, S.Haywood

Rutherford Appleton Laboratory
Chilton DIDCOT
OX11 0QX, UK

e-mails: A.Hicheur@rl.ac.uk, S.Haywood@rl.ac.uk

ATL-INDET-PUB-2005-002
11 October 2005



Abstract

The fundamental linear expansion formalism proposed for the offline alignment of the ATLAS silicon tracking system is presented.

Contents

1	Introduction	1
2	The Least Squares Linear Expansion	1
3	Fitting a Common Event Vertex	3
4	Adding External Constraints	4
5	Derivatives of the residuals	6
5.1	General formalism	6
5.2	Derivatives with respect to the track parameters	7
5.3	Derivatives with respect to the alignment parameters	7
5.4	Derivatives with respect to the common vertex coordinates	8
5.5	Derivatives with respect to the scattering angles	8
5.6	Comments about the derivatives	9
6	Conclusions	10
A	The reference solution for the most general case including the common vertex fit per event and various external constraints	12
B	Two alternative track trajectory parameterisations - without and with the explicit fit of MCS deflection angles	14
C	Some comments on the linear approximation	15
D	An approximate solution	16

1 Introduction

Alignment of the ever growing tracking systems of the modern HEP experiments becomes more and more challenging. Not only do numbers of degrees of freedom rapidly grow but the requirements for the final alignment accuracy get more ambitious due to the desire for greater measurement precision.

The silicon tracking system of the ATLAS detector consists of two subsystems: the PIXel detector and the SCT silicon strip detector. The subsystems are composed of 1744 and 4088 individual modules respectively. The entire system involves nearly 35,000 DoF's! Additional difficulty arises from the fact that the magnetic field in the corresponding tracking volume will be highly nonuniform making all tracking-related calculations (e.g. track extrapolations) nonanalytic and therefore complicated. The basic assessment of the alignment challenge together with physics motivated requirements were formulated in [1] and [2].

In this paper we present the fundamental algebraic formalism at the heart of the current base-line alignment algorithms proposed for the ATLAS silicon. Its main advantage is simplicity and elegance while the major drawback might turn out to be the natural limitation of the linear approximation adopted here and the difficulty of the numerical calculations.

2 The Least Squares Linear Expansion

The alignment algorithm is based on the minimization of a “giant χ^2 ” defined as:

$$\chi^2 = \sum_{\text{tracks}} r^T V^{-1} r \quad \text{where} \quad r_i \equiv (\vec{m}_i - \vec{e}_i(\pi, a)) \cdot \hat{k} \quad (1)$$

Here \vec{e}_i denotes the i 'th intersection point of the extrapolated track with a sensor plane and \vec{m}_i is the position of the associated detector hit. \hat{k} is the unit vector defining the measurement direction for the sensor plane¹. The intersection point depends on both track parameters (π) as well as on the subset of alignment parameters related to the intersected module (a). r is the vector of residuals for a given track. V is the covariance matrix of r . Generally V is not diagonal due to contribution from the Multiple Coulomb Scattering ($V = V_{\text{meas}} + V_{\text{MCS}}$). The notation is, however, equally valid for the case when a track is defined not only by its five perigee parameters but also by the set of MCS deflection angles on the measurement planes. See Appendix B for the details.

We are looking for the solution of the minimisation problem with respect to the alignment parameters. Therefore, we require that the minimum satisfies:

$$\frac{d\chi^2}{da} = 0 \quad \implies \quad \sum_{\text{tracks}} \frac{dr^T}{da} V^{-1} r = 0 \quad (2)$$

As a first step, we need to find the solution for track parameters for any arbitrary alignment. This is nothing else but an ordinary track fit:

$$\frac{\partial \chi^2}{\partial \pi} = 0 \quad \implies \quad \frac{\partial r^T}{\partial \pi} V^{-1} r = 0 \quad (3)$$

¹Recall that for a PIXel detector, each physical hit corresponds to two distinct measurements (2D) and therefore gives rise to two residuals along two measurement directions.

To find the solution we must expand the residuals around their initial values r_0 . We use the first linear term only under the assumption: $\frac{\partial^2 r}{\partial \pi_i \partial \pi_j} \rightarrow 0$, i.e the extrapolated points are a linear function of track parameters, for small changes (see Appendix C for more detailed discussion of the validity of the above assumption). Having expanded around the initial value π_0 , we have:

$$\frac{\partial r^T}{\partial \pi_0} V^{-1} r(\pi_0, a) + \frac{\partial r}{\partial \pi_0} \delta \pi = 0 \quad (4)$$

where $\frac{\partial r}{\partial \pi_0} \equiv \frac{\partial r}{\partial \pi} |_{\pi=\pi_0}$. We then get the solution for the track parameters:

$$\pi = \pi_0 + \delta \pi = \pi_0 - \left(\frac{\partial r^T}{\partial \pi_0} V^{-1} \frac{\partial r}{\partial \pi_0} \right)^{-1} \frac{\partial r^T}{\partial \pi_0} V^{-1} r(\pi_0, a) \quad (5)$$

Effectively, this provides a solution for π as a function of a , which can be reinserted implicitly into the global χ^2 definition. Now let us get back to Equation 2. In exact analogy to the solution for track parameters, we can write down the set of linear equations defining the solution for the alignment:

$$\left(\sum_{tracks} \frac{dr^T}{da_0} V^{-1} \frac{dr}{da_0} \right) \delta a + \sum_{tracks} \frac{dr^T}{da_0} V^{-1} r(\pi_0, a_0) = 0 \quad (6)$$

Because r is a function of π and a the full derivative from Eq. 6 can be written as:

$$\frac{dr}{da_0} = \frac{\partial r}{\partial a_0} + \frac{\partial r}{\partial \pi_0} \frac{d\pi}{da_0} \quad (7)$$

$\frac{d\pi}{da_0}$ is obtained by differentiating Equation 5.

From Eq. 5 and Eq. 7 we get the final expression for the full derivative:

$$\frac{dr}{da_0} = (\mathbf{1} - \mathbf{E}(\mathbf{E}^T V^{-1} \mathbf{E})^{-1} \mathbf{E}^T V^{-1}) \frac{\partial r}{\partial a_0} \quad \text{where} \quad \mathbf{E} \equiv \frac{\partial r}{\partial \pi_0}, \quad (8)$$

which together with Equation 6 gives us the corrections to the alignment parameters which satisfy Eq. 2:

$$\delta a = - \underbrace{\left(\sum_{tracks} \frac{\partial r^T}{\partial a_0} W \frac{\partial r}{\partial a_0} \right)^{-1}}_{\mathcal{M}} \underbrace{\sum_{tracks} \frac{\partial r^T}{\partial a_0} W r(\pi_0, a_0)}_{\mathcal{V}} \quad (9)$$

where we substituted:

$$W \equiv V^{-1} \hat{W} \equiv V^{-1} - V^{-1} \mathbf{E}(\mathbf{E}^T V^{-1} \mathbf{E})^{-1} \mathbf{E}^T V^{-1} \quad (10)$$

and use the relation:

$$\hat{W} V^{-1} \hat{W} = W. \quad (11)$$

$\frac{\partial r}{\partial a_0} \equiv \frac{\partial r}{\partial a} |_{\pi=\pi_0, a=a_0}$. $\mathcal{M} \equiv \left(\sum_{tracks} \frac{\partial r^T}{\partial a_0} W \frac{\partial r}{\partial a_0} \right)$ is a symmetric $n \times n$ matrix where n is the number of alignment DoF's. At this point, one expects that many elements of matrix \mathcal{M} will be null since there is no connection between modules which are well separated in η and ϕ .

3 Fitting a Common Event Vertex

In the previous section we presented the basic calculation leading to the Least Squares solution for alignment parameters with an implicit refit of all considered tracks. A natural and very powerful extension to this formalism comes from the requirement of a common origin (vertex) for certain subsamples of tracks. For simplicity hereafter, they will be denoted as belonging to a common “event”. As before we start from the expression for the global χ^2 (equation 1) but now we have double summation over the events and tracks within an event. Tracks are no more described by their usual five perigee parameters. Only three independent parameters per track remain ($\pi = (\phi, \cot\theta, Q/p_T)$). Impact parameters are replaced by the common vertex for the event ($b = (x_0, y_0, z_0)$). The residuals take the form:

$$r \equiv (\vec{m} - \vec{e}(\pi, b, a)) \cdot \hat{k} \quad (12)$$

and depend explicitly on three different sets of parameters:

1. π – individual track parameters,
2. b – vertex position common for an event,
3. a – alignment parameters common to all events in the dataset.

Let us first find the position of the vertex itself given an arbitrary but fixed alignment. In exact analogy to the solution for a in section 2 we obtain:

$$b = b_0 + \delta b = b_0 - \underbrace{\left(\sum_{tk}^{ev} F^T W F \right)}_{M_b}^{-1} \sum_{tk}^{ev} F^T W r(\pi_0, b_0, a) \quad \text{with} \quad F \equiv \frac{\partial r}{\partial b} \quad (13)$$

It should be noted that in order for this formalism to work, all tracks used within an event must originate from the same b_0 .

The generic form of the solution for alignment parameters resulting from Eq. 12 is:

$$\delta a = - \left(\sum_{ev} \sum_{tk} \frac{dr}{da_0}^T V^{-1} \frac{dr}{da_0} \right)^{-1} \left(\sum_{ev} \sum_{tk} \frac{dr}{da_0}^T V^{-1} r(\pi_0, b_0, a_0) \right) \quad (14)$$

but now $\frac{dr}{da}$ has more complicated structure.

$$\begin{aligned} \frac{dr}{da} &= \frac{\partial r}{\partial a} + E \frac{d\pi}{da} + F \frac{db}{da} \\ \frac{d\pi}{da} &= - (E^T V^{-1} E)^{-1} E^T V^{-1} \left(\frac{\partial r}{\partial a} + F \frac{db}{da} \right) \\ \frac{db}{da} &= - M_b^{-1} \left(\sum_{tk}^{ev} F^T W \frac{\partial r}{\partial a} \right) \end{aligned} \quad (15)$$

where E, F and W are defined as before. From Eq. 15 we further find:

$$\frac{dr}{da} = \hat{W} \alpha_a^r \quad \text{where} \quad \alpha_a^r \equiv \frac{\partial r}{\partial a} - F M_b^{-1} \left(\sum_{tk}^{ev} F^T W \frac{\partial r}{\partial a} \right) \quad (16)$$

Equation 14 and Eq. 16 combined yield the solution for the alignment:

$$\delta a = - \underbrace{\left(\sum_{ev} \sum_{tk} \alpha_a^{rT} W \alpha_a^r \right)}_{\mathcal{M}}^{-1} \underbrace{\sum_{ev} \sum_{tk} \alpha_a^{rT} W r(\pi_0, b_0, a_0)}_{\mathcal{V}} \quad (17)$$

where we used again the identity 11. Note that with the imposed common vertex, the solution again includes the $n \times n$ symmetric error matrix \mathcal{M}^{-1} . However, this time, none of the matrix elements are identically zero. Given a large number of tracks, the matrix will be fully populated.

Correlations have three sources:

1. DoF's of a single module,
2. correlations through a common track,
3. tracks emerging from a common vertex!

4 Adding External Constraints

The solutions presented so far do not allow for any external constraints imposed on the solution. In the most general case we could have constraints on:

1. track parameters (π),
2. beam constraint (b),
3. alignment parameters (a).

The above should be thought of as additional bi-linear terms in the χ^2 expression. The constraints in general may not be linear but these have to be linearised before they can enter the proposed formalism. In the following we sketch an example of a constraint on track parameters within the formalism without the vertex fit. To start with, let us recall the implication of the constraints on the individual track fit.

$$\chi^2 = r(\pi)^T V^{-1} r(\pi) + (\pi - x)^T S^{-1} (\pi - x), \quad (18)$$

where vector x and covariance matrix S define the constraint of π . The linear expansion of the χ^2 expression 18 leads to the solution:

$$\delta\pi = - (\mathbf{E}^T V^{-1} \mathbf{E} + S^{-1})^{-1} (\mathbf{E}^T V^{-1} r(\pi_0, a) + S^{-1} (\pi_0 - x)) \quad (19)$$

Propagation of the constraint to the global alignment fit does not present any major complication. Let us start from the modified χ^2 definition:

$$\chi^2 = \sum_{tracks} (r(\pi, a)^T V^{-1} r(\pi, a) + (\pi - x)^T S^{-1} (\pi - x)), \quad (20)$$

The following lines show briefly how the final solution is obtained using the first order linear expansion of the χ^2 expression around the minimum. For brevity we adopt the simplified notation: $J \equiv \mathbf{E}^T V^{-1} \mathbf{E} + S^{-1}$, $r_0 \equiv r(\pi_0, a_0)$ & $\frac{\partial r}{\partial a} \equiv \frac{\partial r}{\partial a_0}$ in all cases.

$$\begin{aligned} \frac{d\chi^2}{da} &= 0 \\ \Downarrow \\ \sum_{tracks} \left(\frac{dr}{da}^T V^{-1} (r_0 + \frac{dr}{da} \delta a) + \frac{d\pi}{da}^T S^{-1} (\pi_0 + \frac{d\pi}{da} \delta a - x) \right) &= 0 \\ \frac{dr}{da} &= \frac{\partial r}{\partial a} + \mathbf{E} \frac{d\pi}{da} \quad \text{and} \quad \frac{d\pi}{da} = -J^{-1} \mathbf{E}^T V^{-1} \frac{\partial r}{\partial a} \end{aligned} \quad (21)$$

$$\begin{aligned}
& \sum_{tracks} \frac{\partial r^T}{\partial a} \left(\hat{W}^T V^{-1} \hat{W} + V^{-1} E J^{-1} S^{-1} J^{-1} E^T V^{-1} \right) \frac{\partial r}{\partial a} \delta a \\
& + \sum_{tracks} \left(\frac{\partial r^T}{\partial a} \hat{W}^T V^{-1} r_0 \right) \\
& - \sum_{tracks} \left(\frac{\partial r^T}{\partial a} V^{-1} E J^{-1} S^{-1} (\pi_0 - x) \right) = 0
\end{aligned} \tag{22}$$

where we use the redefined expression for W :

$$\hat{W} = \mathbf{1} - E J^{-1} E^T V^{-1} \quad \text{and} \quad W \equiv V^{-1} \hat{W} \tag{23}$$

Since

$$\hat{W}^T V^{-1} \hat{W} + V^{-1} E J^{-1} S^{-1} J^{-1} E^T V^{-1} = W \tag{24}$$

we finally get:

$$\begin{aligned}
\delta a = - & \overbrace{\left(\sum_{tracks} \frac{\partial r^T}{\partial a} W \frac{\partial r}{\partial a} \right)^{-1}}^{\mathcal{M}^{-1}} \\
& \left(\sum_{tracks} \frac{\partial r^T}{\partial a} W r_0 \right. \\
& \left. - \sum_{tracks} \frac{\partial r^T}{\partial a} V^{-1} E J^{-1} S^{-1} (\pi_0 - x) \right)
\end{aligned} \tag{25}$$

The above formalism may be useful for implementing a beam spot constraint, constraining track parameters from external tracking systems, for example calorimetry (e.g. E/p) or eventually building the constraints based on reconstructed physics events (mass constraint from narrow resonances, etc.). We do not have space here to derive the formulae for the most general case. The main results and generalised solution are given in the form of a “reference card” in Appendix A.

5 Derivatives of the residuals

In the global χ^2 minimisation detailed in the previous sections, one can see that the main ingredients to derive the alignment parameters are error matrices and the derivatives of residuals with regard to track and alignment parameters, vertex coordinates, and scattering angles. The error matrices are inputs from measurements or fits performed earlier in the reconstruction chain (e.g scattering angles RMS). The aim of this section is to explain how the different derivatives of the residuals are obtained.

5.1 General formalism

Consider $\hat{X}, \hat{Y}, \hat{Z}$ -the unit vectors defining the module local frame. \hat{X} is in the direction crossing the strips and \hat{Y} is along the strips. \hat{Z} is the normal to the plane.

The two residuals, in the detector plane, are then defined by:

$$\begin{aligned} res_X &= (\vec{m} - \vec{e}) \cdot \hat{X} \\ res_Y &= (\vec{m} - \vec{e}) \cdot \hat{Y} \end{aligned} \quad (26)$$

The derivatives of these quantities wrt a given parameter p are:

$$\begin{aligned} \frac{\partial res_X}{\partial p} &= -\frac{d\vec{e}}{dp} \cdot \hat{X} \\ \frac{\partial res_Y}{\partial p} &= -\frac{d\vec{e}}{dp} \cdot \hat{Y} \end{aligned} \quad (27)$$

The position of a point on the trajectory is characterised by l , the path length. The intersection point is defined by $(\vec{e} - \vec{m}) \cdot \hat{Z} = 0$. Usually, this equation is solved by iteration, as it cannot be solved analytically.

From this equation, we extract l_{int} , i.e the path length at the intersection point. Note that this quantity is an implicit function of p .

Now we can compute $\frac{d\vec{e}}{dp}$:

$$\frac{d\vec{e}}{dp} = \frac{\partial \vec{e}}{\partial p} + \frac{\partial \vec{e}}{\partial l_{int}} \cdot \frac{dl_{int}}{dp} \quad (28)$$

where $\frac{\partial \vec{e}}{\partial l_{int}} \equiv \frac{\partial \vec{e}}{\partial l} |_{l=l_{int}}$ and $\frac{dl_{int}}{dp}$ is derived by differentiating the intersection equation, $(\vec{e} - \vec{m}) \cdot \hat{Z} = 0$, with respect to p and taking into account the fact that l_{int} is an implicit function of p :

$$\frac{d((\vec{e} - \vec{m}) \cdot \hat{Z})}{dp} = \frac{d(\vec{e} \cdot \hat{Z})}{dp} = \frac{\partial(\vec{e} \cdot \hat{Z})}{\partial p} + \frac{\partial(\vec{e} \cdot \hat{Z})}{\partial l_{int}} \frac{dl_{int}}{dp} = 0$$

And then:

$$\frac{dl_{int}}{dp} = -\frac{\frac{\partial(\vec{e} \cdot \hat{Z})}{\partial p}}{\frac{\partial(\vec{e} \cdot \hat{Z})}{\partial l_{int}}} \quad (29)$$

Then, the final formulae for the residual derivatives are:

$$\begin{aligned}\frac{\partial res_X}{\partial p} &= -\left(\frac{\partial \vec{e}}{\partial p} - \frac{\partial \vec{e}}{\partial l_{int}} \frac{\partial(\vec{e} \cdot \hat{Z})}{\partial l_{int}}\right) \cdot \hat{X} \\ \frac{\partial res_Y}{\partial p} &= -\left(\frac{\partial \vec{e}}{\partial p} - \frac{\partial \vec{e}}{\partial l_{int}} \frac{\partial(\vec{e} \cdot \hat{Z})}{\partial l_{int}}\right) \cdot \hat{Y}\end{aligned}\quad (30)$$

which can be rewritten:

$$\frac{\partial res_\alpha}{\partial p} = -\frac{\partial \vec{e}}{\partial p} \cdot \left(\hat{\alpha} - \frac{\partial \vec{e}}{\partial l_{int}} \cdot \hat{\alpha} \cdot \hat{Z}\right) \quad (31)$$

where $\alpha = X, Y$.

$\frac{\partial \vec{e}}{\partial l_{int}}$ is the tangent to the track at the intersection point: in other words, it gives the direction of the straight line representing the local approximation of the track.

In the local module frame, we define the direction of the track as (Φ, Θ) . Then, the above equation becomes:

$$\begin{aligned}\frac{\partial res_X}{\partial p} &= -\frac{\partial \vec{e}}{\partial p} \cdot (\hat{X} - \cos(\Phi) \tan(\Theta) \hat{Z}) \\ \frac{\partial res_Y}{\partial p} &= -\frac{\partial \vec{e}}{\partial p} \cdot (\hat{Y} - \sin(\Phi) \tan(\Theta) \hat{Z})\end{aligned}\quad (32)$$

The vectors $\overrightarrow{proj}_X \equiv \hat{X} - \cos(\Phi) \tan(\Theta) \hat{Z}$ and $\overrightarrow{proj}_Y \equiv \hat{Y} - \sin(\Phi) \tan(\Theta) \hat{Z}$ are the projection directions along which the change of the extrapolation is computed.

5.2 Derivatives with respect to the track parameters

The track parameters are denoted by $\pi = \{a_0, z_0, \phi_0, \cotan(\theta), q/p_T\}$. From the results found before, the derivatives of the residuals can be written as:

$$\frac{\partial res_\alpha}{\partial \pi} = -\frac{\partial \vec{e}}{\partial \pi} \cdot \overrightarrow{proj}_\alpha \quad (33)$$

where $\alpha = X, Y$.

5.3 Derivatives with respect to the alignment parameters

The vision we adopt here is that the hit is “stuck” to the module and we study the effect of the translations and rotations of the module on the intersection with the extrapolated track, and therefore on the residuals. Note that here, we are defining partial derivatives that only involve residuals of hits belonging to the module being displaced. Generally, the motion of a given module can have an influence on residuals calculated in other modules and this is taken into account in the total derivative of the residuals, $\frac{dr}{da}$, defined in the equations 7 and 15.

Let's call T_X, T_Y, T_Z the translation parameters and R_X, R_Y, R_Z the rotation parameters.

Following the same methodology developed in the previous subsection, we find:

$$\begin{aligned}\frac{\partial res_\beta}{\partial T_\alpha} &= \hat{\alpha} \cdot \overrightarrow{proj}_\beta \\ \frac{\partial res_\beta}{\partial R_\alpha} &= (\hat{\alpha} \times \overrightarrow{lever}) \cdot \overrightarrow{proj}_\beta\end{aligned}\quad (34)$$

where $\alpha = X, Y, Z$; $\beta = X, Y$, $\overrightarrow{lever} = \overrightarrow{OM}$, O being the centre of the module and M is the intersection point of the track.

5.4 Derivatives with respect to the common vertex coordinates

Let (v_x, v_y, v_z) be the primary vertex position in the global frame.

The derivatives of the residuals with regard to the primary vertex coordinates are therefore:

$$\frac{\partial res_\alpha}{\partial v_i} = \hat{n}_i \cdot \overrightarrow{proj}_\alpha \quad (35)$$

where $\alpha = X, Y$, $i = x, y, z$, and $\hat{n}_x = (1, 0, 0)$, $\hat{n}_y = (0, 1, 0)$, $\hat{n}_z = (0, 0, 1)$, in the global frame.

5.5 Derivatives with respect to the scattering angles

Consider two scattering planes (i.e modules) a and b . We study the effect of the scattering occuring through plane a on the residuals computed in plane b .

If $a \geq b$, i.e the track passes through plane b before plane a , then the answer is trivial: the derivatives of the residuals computed in plane b with regard to the scattering occuring later in plane a are null.

If $a < b$: following the same formalism used in the previous paragraphs, we get:

$$\frac{\partial res_\alpha^b}{\partial \Psi^a} = -\frac{\partial \vec{e}^b}{\partial \Psi^a} \cdot \overrightarrow{proj}_\alpha^b \quad (36)$$

where $\Psi = \Theta, \Phi$ are the scattering angles.

The quantity $\frac{\partial \vec{e}^b}{\partial \Psi^a}$ represents the rate of the variation $\delta \vec{e}^b$ of the intersection point of the track with plane b , given a change in the scattering angle in plane a , $\delta \Psi^a$ (the path length is fixed here).

Changing the angle Θ^a by $\delta \Theta^a$ is equivalent to a rotation of angle $\delta \Theta^a$ about the axis \vec{u}_ϕ^a whose centre is \vec{e}^a , where $(\vec{u}_r^a, \vec{u}_\phi^a, \vec{u}_\theta^a)$ are the local axes corresponding to the local frame of module a , in spherical coordinates (\vec{u}_r^a is along the track direction at the point \vec{e}^a). Thus, the variation of the position of the intersection point in plane b is:

$$\begin{aligned}\delta \vec{e}^b &= \delta \Theta^a \vec{u}_\phi^a \times (\vec{e}^b - \vec{e}^a) \\ &= \delta \Theta^a (-\sin(\Phi^a) \hat{X}^a + \cos(\Phi^a) \hat{Y}^a) \times (\vec{e}^b - \vec{e}^a)\end{aligned}\quad (37)$$

In the same way, changing the angle Φ^a by $\delta\Phi^a$ is equivalent to a rotation of angle $\delta\Phi^a$ about the axis \hat{Z}^a whose centre is \bar{e}^a :

$$\delta\bar{e}^b = \delta\Phi^a \hat{Z}^a \times (\bar{e}^b - \bar{e}^a) \quad (38)$$

Equation 36 becomes then:

$$\begin{aligned} \frac{\partial res_\alpha^b}{\partial \Theta^a} &= -[(-\sin(\Phi^a)\hat{X}^a + \cos(\Phi^a)\hat{Y}^a) \times (\bar{e}^b - \bar{e}^a)].\overrightarrow{proj}_\alpha^b \\ \frac{\partial res_\alpha^b}{\partial \Phi^a} &= -[\hat{Z}^a \times (\bar{e}^b - \bar{e}^a)].\overrightarrow{proj}_\alpha^b \end{aligned} \quad (39)$$

5.6 Comments about the derivatives

A certain number of points are worth being mentionned about these derivatives:

- The formulae shown in equation 32 are general and do not depend on any assumption about the trajectory.
- However, for the derivatives with regard to track parameters, one needs underlying assumptions, such as constant magnetic field, to get them analytically. The possible concern is that the non-uniformity of the magnetic field will have to be taken into account, particularly in the endcaps region where the field varies more rapidly.
- Following the comments in the previous point, the alternative we envisage is that the derivatives with respect to the track parameters may be calculated numerically or semi-numerically since from equation 32, only $\frac{\partial \vec{e}}{\partial p}$ would have to be calculated numerically. We can hope that the analytical derivatives will be precise enough, at least for the barrel region, and we will try to avoid, whenever it is possible, to use a numerical derivation, to save computing time.

6 Conclusions

A simple, generic formalism to solve an arbitrary track alignment problem has been presented. The solution is based on the minimisation of the generalised χ^2 and uses the first order linear expansion of the latter. The formalism has number of apparent features: Good points:

- It makes maximal use of the reconstruction information, in a well defined way allowing all the DoF to be fitted (apart from few degenerate modes).
- In principal, it converges in one pass (provided that the update required to the alignment is not too big)
- It can deal with the non-uniform B-field.
- Possibility of adding constraints: E/p comes as a natural constraint, adding the following term to the χ^2 : $(E/p - 1)^2/\sigma^2$, as well as $Z \rightarrow \mu^+\mu^-$ for the alignment-dependent mass scale calibration, adding $(m_{\mu\mu} - m_Z)^2/\Gamma^2$ to the global χ^2 .

Points of concern:

- The solution requires solving a system of n linear equations with n parameters where n is the number of DoF's of the system. The task is numerically challenging. Some considerations about possible strategies are given in Appendix D.
- First order linear expansion to fit tracks/alignment parameters, valid for small distortions only - may require iterations.

References

- [1] S. Haywood, *Offline Alignment and Calibration of the Inner Detector*, ATLAS Internal Note, ATL-COM-INDET-99-001 (1999)
- [2] ATLAS Collaboration, *ATLAS Detector and Physics Performance TDR*, ATLAS TDR 15, CERN/LHCC/99-15, 25 May 1999.

Appendices

A The reference solution for the most general case including the common vertex fit per event and various external constraints

In the following page, we present the key expressions defining the solution in the most general case. The colour coding should help the reader associate various bits of the final expression with their origin. We have:

- Constraints on track parameters (π),
- Contributions from beam constraint,
- Constraints on alignment parameters (a),
- Fit for the vertex position (b).

For brevity we use the simplified notation:

$$\frac{\partial r}{\partial \pi} \equiv \frac{\partial r}{\partial \pi} |_{\pi=\pi_0, b=b_0, a=a_0}, \quad \frac{\partial r}{\partial b} \equiv \frac{\partial r}{\partial b} |_{\pi=\pi_0, b=b_0, a=a_0}, \quad \frac{\partial r}{\partial a} \equiv \frac{\partial r}{\partial a} |_{\pi=\pi_0, b=b_0, a=a_0} \quad \text{in all cases.}$$

**The reference solution for the most general case
including the common vertex fit per event and external constraints**

$$\chi^2 = \sum_{ev} \left(\sum_{tk}^{ev} (r(\pi, b, a)^T V^{-1} r(\pi, b, a) + (\pi - x)^T S^{-1} (\pi - x)) + (b - v)^T Q^{-1} (b - v) \right) + (a - t)^T U^{-1} (a - t) \quad (1)$$

$$\delta\pi = -J^{-1} (E^T V^{-1} r(\pi_0, b, a) + S^{-1} (\pi_0 - x)) \quad (2)$$

$$\delta b = - \underbrace{\left(\sum_{tk}^{ev} F^T W F + Q^{-1} \right)^{-1}}_{M_b^{-1}} \left(\sum_{tk}^{ev} (F^T W r(\pi_0, b_0, a) - F^T V^{-1} E J^{-1} S^{-1} (\pi_0 - x)) + Q^{-1} (b_0 - v) \right) \quad (3)$$

where:

$$E \equiv \frac{\partial r}{\partial \pi}, \quad F \equiv \frac{\partial r}{\partial b}, \quad J \equiv E^T V^{-1} E + S^{-1}, \quad W = V^{-1} - V^{-1} E J^{-1} E^T V^{-1} \quad (4)$$

Defining:

$$\alpha_a^r \equiv \frac{\partial r}{\partial a} + F \frac{db}{da} = \frac{\partial r}{\partial a} - F M_b^{-1} \left(\sum_{tk}^{ev} F^T W \frac{\partial r}{\partial a} \right) \quad (5)$$

We can write the final solution for the alignment as:

$$\delta a = - \overbrace{\left\{ \sum_{ev} \left(\sum_{tk}^{ev} \alpha_a^{rT} W \alpha_a^r + \frac{db^T}{da} Q^{-1} \frac{db}{da} \right) + U^{-1} \right\}^{-1}}^{\mathcal{M}^{-1}} \left\{ \sum_{ev} \left(\sum_{tk}^{ev} \left(\alpha_a^{rT} W r(\pi_0, b_0, a_0) + \frac{d\pi}{da}^T S^{-1} (\pi_0 - x) \right) + \frac{db^T}{da} Q^{-1} (b_0 - v) \right) + U^{-1} (a_0 - t) \right\} \quad (6)$$

where from (2), (3) and (5) we have:

$$\frac{d\pi}{da} = -J^{-1} E^T V^{-1} \alpha_a^r \quad \text{and} \quad \frac{db}{da} = -M_b^{-1} \sum_{tk}^{ev} F^T W \frac{\partial r}{\partial a} \quad (7)$$

B Two alternative track trajectory parameterisations - without and with the explicit fit of MCS deflection angles

In the “standard” representation, the track is described by five parameters (position and direction at the perigee). Hits associated to the track have residuals resulting from propagating the track through the magnetic field volume in the absence of scattering material. Therefore, the error matrix of the residuals is composed of a diagonal matrix corresponding to the measurement errors and a non-diagonal matrix responsible for the Multiple Coulomb Scattering on the measurement surfaces.

$$V = V_{\text{meas}} + V_{\text{MCS}} \quad \text{where} \quad V_{\text{MCS}} = \frac{\partial r}{\partial \alpha} N_{\text{MCS}} \frac{\partial r^T}{\partial \alpha} \quad (1)$$

N_{MCS} is a diagonal matrix describing the rms of expected deflection angle on a given scattering plane (according to the theory of Molière) and $\partial r/\partial \alpha$ are derivatives relating the deflection angle (α) to the change of track extrapolation to the consecutive measurement planes. V_{meas} describes the intrinsic resolution of the sensitive devices and is diagonal by construction.

An alternative approach defines a track not only by the five parameters of its helix but also by a sequence of scattering angles attributed to every scattering plane it traverses. In this approach, we are bound to identify scattering planes with the actual measurement planes. In this formalism some objects change their identity:

$$r \longrightarrow r' \equiv \begin{pmatrix} r' \\ \alpha \end{pmatrix}, \quad \pi \longrightarrow \pi \equiv \begin{pmatrix} \pi \\ \alpha \end{pmatrix} \quad \text{and} \quad V \longrightarrow V' \equiv \begin{pmatrix} V_{\text{meas}} & 0 \\ 0 & N_{\text{MCS}} \end{pmatrix}. \quad (2)$$

At the same time the apparent residuals have to be redefined to accomodate the zig-zagging track trajectory:

$$r' = r + \frac{\partial r}{\partial \alpha} \alpha, \quad (3)$$

as well as the matrix of hit residual derivatives with respect to track parameters:

$$\frac{\partial r'}{\partial \pi} = \begin{pmatrix} \frac{\partial r}{\partial \pi} & \frac{\partial r}{\partial \alpha} \\ 0 & \mathbf{1} \end{pmatrix}. \quad (4)$$

It can be shown that track fits using the two above formalisms lead to the same track perigee parameters. The proof is beyond the scope of this paper²

As can be seen from Eqs. (1) through (4) the procedure is straightforward and requires only two computations:

$$V = V_{\text{meas}} + \frac{\partial r}{\partial \alpha} N_{\text{MCS}} \frac{\partial r^T}{\partial \alpha} \quad \text{and} \quad r = r' - \frac{\partial r}{\partial \alpha} \alpha \quad (5)$$

It has been verified that the fit results do not change after the above compactification.

²http://hepwww.rl.ac.uk/Atlas-SCT/alignment/algorithm_ms.pdf

C Some comments on the linear approximation

To solve for the alignment, we minimize the χ^2 which, in the absence of explicit constraints, is given by the residuals (and deflection angles) and their weight matrix:

$$\chi^2 = \sum_{\text{tracks}} r^T V^{-1} r \quad \text{and} \quad r = m - e(\pi, a) \quad (1)$$

The usual way of solving analytically χ^2 problems is to assume that the quantities used to build the χ^2 (here the residuals) are linear in the parameters of interest (here the alignment parameters). This leads to the following parabolic expansion near the minimum:

$$\chi^2 = \chi_0^2 + \frac{d\chi^2}{da} \cdot \delta a + \frac{1}{2} \delta a^T \cdot \frac{d^2\chi^2}{dada} \cdot \delta a \quad (2)$$

where $\chi_0^2 = \chi^2(a_0)$.

The minimization leads to the following estimation of the alignment parameters:

$$\delta a = -\left(\frac{d^2\chi^2}{dada}\right)^{-1} \left(\frac{d\chi^2}{da}\right) \quad (3)$$

The complete expression of an element of matrix $\frac{d^2\chi^2}{dada}$ is

$$\frac{d^2\chi^2}{da_i da_j} = \sum_{\text{tracks}} \underbrace{\frac{dr^T}{da_i} V^{-1} \frac{dr}{da_j}}_{\text{first term}} + \underbrace{\frac{d^2 r^T}{da_i da_j} V^{-1} r}_{\text{second term}} \quad (4)$$

According to the linear approximation, the second term is dropped. It turns out that in some cases, for particular tracks, the first term above is much smaller than the term which is dropped. However, this is not a problem since the expression involves a sum over many tracks and so, even if the first term for a given track vanishes, it is likely that its second term is much smaller than contributions from other tracks.

Furthermore, for the misalignments expected, it is anticipated that the linear approximation will be good. Even if some of the initial misalignments are large, by applying the alignment procedure once, any large uncertainties in module position will be significantly reduced, so that changes to the parameters on a second iteration will be compatible with the assumption of linearity to the required precision.

Having said this, there may be some consequences of these approximations and these will be examined in subsequent studies on the performance of the alignment algorithm.

D An approximate solution

The final matrix \mathcal{M} is generally singular. The null eigenvalues correspond to the global DoF's of the entire system. Moreover, the remaining eigen spectrum spans over 10 orders of magnitude making the solution numerically challenging. The simplest solution by substitution is not good here.

An elegant alternative is provided by diagonalisation of the matrix, $\mathcal{M}_D = U\mathcal{M}U^T$. The operation is, however, quite CPU time consuming (6× more than simple matrix inversion!).

The workaround might come from the iterative procedures using approximate matrices. In the simplest case, we define an “ideal” matrix $\hat{\mathcal{M}}$,³ invert it and use it to obtain the solution by iteration:

$$a = - \sum_{n=1}^{\infty} \left\{ 1 - \hat{\mathcal{M}}^{-1}\mathcal{M} \right\}^n \hat{\mathcal{M}}^{-1} \frac{dr^T}{da} V^{-1} r_0, \quad (1)$$

which converges to $-\mathcal{M}^{-1} \frac{dr^T}{da} V^{-1} r_0$ if $|1 - \hat{\mathcal{M}}^{-1}\mathcal{M}| < 1$.

³The matrix can be obtained from simulated events using the nominal detector geometry.